PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 14/435, 14/705, C12N 5/10, 15/10, 15/12, 15/63

A1

(11) International Publication Number: WO 99/29725

(43) International Publication Date: 17 June 1999 (17.06.99)

(21) International Application Number: PCT/US98/26422

(22) International Filing Date: 11 December 1998 (11.12.98)

(30) Priority Data:
60/069,379 12 December 1997 (12.12.97) US

(71) Applicant *(for all designated States except US)*: MERCK & CO., INC. [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: CHEN, Fang [US/US]; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(74) Common Representative: MERCK & CO., INC.; 126 East Lincoln Avenue, Rahway, NJ 07065 (US).

(81) Designated States: CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: DNA MOLECULES ENCODING HUMAN NUCLEAR RECEPTOR PROTEIN, nNR5

(57) Abstract

The present invention discloses the isolation and characterization of cDNA molecules encoding a novel member to the human nuclear receptor superfamily, designated nNR5. Also within the scope of the disclosure are recombinant vectors, recombinant host cells, methods of screening for modulators of nNR5 activity, and production of antibodies against nNR5, or epitopes thereof.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

TITLE OF THE INVENTION
DNA MOLECULES ENCODING HUMAN NUCLEAR RECEPTOR PROTEIN, nNR5

FIELD OF THE INVENTION

The present invention relates in part to isolated nucleic acid molecules (polynucleotides) which encode vertebrate nuclear receptor proteins, and especially human nuclear receptor proteins as exemplified throughout this specification as nNR5. The present invention also relates to recombinant vectors and recombinant hosts which contain a DNA fragment encoding nNR5, substantially purified forms of associated human nNR5 protein, human mutant proteins, and methods associated with identifying compounds which modulate nNR5 activity.

BACKGROUND OF THE INVENTION

The nuclear receptor superfamily, which includes steroid hormone receptors, are small chemical ligand-inducible transcription factors which have been shown to play roles in controlling development, differentiation and physiological function. Isolation of cDNA clones encoding nuclear receptors reveal several characteristics. First, the NH_2-terminal regions, which vary in length between receptors, is hypervariable with low homology between family members. There are three internal regions of conservation, referred to as domain I, II and III. Region I is a cysteine-rich region which is referred to as the DNA binding domain (DBD). Regions II and III are within the COOH-

terminal region of the protein and is also referred to as the ligand binding domain (LBD). For a review, see Power et al. (1992, *Trends in Pharmaceutical Sciences* 13: 318-323).

The lipophilic hormones that activate steroid receptors are known to be associated with human diseases. Therefore, the respective nuclear receptors have been identified as possible targets for therapeutic intervention. For a review of the mechanism of action of various steroid hormone receptors, see Tsai and O'Malley (1994, *Annu. Rev. Biochem.* 63: 451-486).

Recent work with non-steroid nuclear receptors has also shown the potential as drug targets for therapeutic intervention. This work reports that peroxisome proliferator activated receptor g (PPARg), identified by a conserved DBD region, promotes adipocyte differentiation upon activation and that thiazolidinediones, a class of antidiabetic drugs, function through PPARg (Tontonoz et al., 1994, *Cell* 79: 1147-1156; Lehmann et al., 1995, *J. Biol. Chem.* 270(22): 12953-12956; Teboul et al., 1995, *J. Biol. Chem.* 270(47): 28183-28187). This indicates that PPARg plays a role in glucose homeostasis and lipid metabolism.

Wang et al. (1989, *Nature* 340: 163-166) show data which prompted the authors to classify the COUP transcription factor (COUP-TF) as a member of the nuclear receptor superfamily.

Mangelsdorf et al. (1995, *Cell* 83: 835-839) provide a review of known members of the nuclear receptor superfamily.

It would be advantageous to identify additional genes which are members of the nuclear receptor superfamily, especially vertebrate members from such species as human, rat and mouse. A nucleic acid molecule expressing a nuclear receptor protein will be useful in screening for compounds acting as a modulator of cell differentiation, cell development and physiological function. The present invention addresses and meets these needs by disclosing isolated nucleic acid molecules which express a human nuclear receptor protein which will have a role in cell differentiation and development.

SUMMARY OF THE INVENTION

The present invention relates to isolated nucleic acid molecules (polynucleotides) which encode novel nuclear receptor proteins which are herein designated as members of the nuclear receptor superfamily. The isolated polynucleotides of the present invention encode vertebrate members of this nuclear receptor superfamily, and preferably human nuclear receptor proteins, such as the human nuclear receptor protein exemplified and referred to throughout this specification as nNR5. The nuclear receptor proteins encoded by the isolated polynucleotides of the present invention are involved in the regulation of *in vivo* cell proliferation and/or cell development.

The present invention also relates to isolated nucleic acid fragments which encode mRNA expressing a biologically active novel vertebrate nuclear receptor which belongs to the nuclear receptor superfamily. A preferred embodiment relates to isolated nucleic acid fragments of SEQ ID NO: 1 which encode mRNA expressing a biologically functional derivative of nNR5. Any such nucleic acid fragment will encode either a protein or protein fragment comprising at least an intracellular DNA-binding domain and/or ligand binding domain, domains conserved throughout the human nuclear receptor family domain which exist in nNR5 (SEQ ID NO:2). Any such polynucleotide includes but is not necessarily limited to nucleotide substitutions, deletions, additions, amino-terminal truncations and carboxy-terminal truncations such that these mutations encode mRNA which express a protein or protein fragment of diagnostic, therapeutic or prophylactic use and would be useful for screening for agonists and/or antagonists of nNR5.

The isolated nucleic acid molecule of the present invention may include a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which may be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention may also include a ribonucleic acid molecule (RNA).

The present invention also relates to recombinant vectors and recombinant hosts, both prokaryotic and eukaryotic, which contain the substantially purified nucleic acid molecules disclosed throughout this specification.

A preferred embodiment of the present invention is an isolated cDNA molecule which encodes a human nuclear receptor protein, wherein said protein is substantially expressed in eye, especially the retina. The isolated cDNA molecules and expressed and isolated nuclear receptor proteins of the present invention are involved in the regulation of gene expression. Due to its high expression in retinal tissue, nNR5 should play an important role in eye function. Therapeutic compounds may be selected which interact with and regulate nNR5 activity in retina tissue which may be involved with diseases of the eye, including but not limited to cataracts and glaucoma, as well as retina-specific diseases such as diabetes mellitus, retinitis pigmentosa, macular degeneration, retinal detachment and retinablastoma.

An especially preferred embodiment of the present invention is disclosed in Figure 1A-B and SEQ ID NO: 1, an isolated human cDNA encoding a novel nuclear trans-acting receptor protein, nNR5.

Another preferred aspect of the present invention relates to a substantially purified form of the novel nuclear trans-acting receptor protein, nNR5, which is disclosed in Figures 2A-B and Figure 3 and as set forth in SEQ ID NO:2.

Another embodiment of the present invention relates to an isolated cDNA molecule encoding nNR5 which also contains a single intron from nucleotide # 971 to nucleotide # 1847 of SEQ ID NO: 18.

The present invention also relates to biologically functional derivatives of nNR5 as set forth as SEQ ID NO:2, including but not limited to nNR5 mutants and biologically active fragments such as amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations, such that these fragments provide for proteins or protein fragments of diagnostic,

therapeutic or prophylactic use and would be useful for screening for agonists and/or antagonists of nNR5 function.

The present invention also relates to polyclonal and monoclonal antibodies raised in response to either the human form of nNR5 disclosed herein, or a biologically functional derivative thereof. It will be especially preferable to raise antibodies against epitopes within the NH_2-terminal domain of nNR5, which show the least homology to other known proteins belonging to the human nuclear receptor superfamily. To this end, the DNA molecules, RNA molecules, recombinant protein and antibodies of the present invention may be used to screen and measure levels of human nNR5. The recombinant proteins, DNA molecules, RNA molecules and antibodies lend themselves to the formulation of kits suitable for the detection and typing of human nNR5.

The present invention also relates to isolated nucleic acid molecules which are fusion constructions expressing fusion proteins useful in assays to identify compounds which modulate wild-type human nNR5 activity. A preferred aspect of this portion of the invention includes, but is not limited to, glutathione S-transferase GST-nNR5 fusion constructs. These fusion constructs include, but are not limited to, all or a portion of the ligand-binding domain of nNR5, respectively, as an in-frame fusion at the carboxy terminus of the GST gene. The disclosure of SEQ ID NOS:1-2 allow the artisan of ordinary skill to construct any such nucleic acid molecule encoding a GST-nuclear receptor fusion protein. Soluble recombinant GST-nuclear receptor fusion proteins may be expressed in various expression systems, including *Spodoptera frugiperda* (Sf21) insect cells (Invitrogen) using a baculovirus expression vector (e.g., Bac-N-Blue DNA from Invitrogen or pAcG2T from Pharmingen).

It is an object of the present invention to provide an isolated nucleic acid molecule which encodes a novel form of a nuclear receptor protein such as human nNR5, human nuclear receptor protein fragments of full length proteins such as nNR5, and mutants which are derivatives of SEQ ID NO:2. Any such polynucleotide includes but is not necessarily limited to nucleotide substitutions, deletions, additions,

amino-terminal truncations and carboxy-terminal truncations such that these mutations encode mRNA which express a protein or protein fragment of diagnostic, therapeutic or prophylactic use and would be useful for screening for agonists and/or antagonists for nNR5 function.

Another object of this invention is tissue typing using probes or antibodies of this invention. In a particular embodiment, polynucleotide probes are used to identify tissues expressing nNR5 mRNA. In another embodiment, probes or antibodies can be used to identify a type of tissue based on nNR5 expression or display of nNR5 receptors.

It is a further object of the present invention to provide the human nuclear receptor proteins or protein fragments encoded by the nucleic acid molecules referred to in the preceding paragraph.

It is a further object of the present invention to provide recombinant vectors and recombinant host cells which comprise a nucleic acid sequence encoding human nNR5 or a biological equivalent thereof.

It is an object of the present invention to provide a substantially purified form of nNR5, as set forth in SEQ ID NO:2.

It is an object of the present invention to provide for biologically functional derivatives of nNR5, including but not necessarily limited to amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations such that these fragment and/or mutants provide for proteins or protein fragments of diagnostic, therapeutic or prophylactic use.

It is also an object of the present invention to provide for nNR5-based in-frame fusion constructions, methods of expressing these fusion constructions and biological equivalents disclosed herein, related assays, recombinant cells expressing these constructs and agonistic and/or antagonistic compounds identified through the use DNA molecules encoding human nuclear receptor proteins such as nNR5 and nNR2.

As used herein, "DBD" refers to DNA binding domain.

As used herein, "LBD" refers to ligand binding domain.

As used herein, the term "mammalian host" refers to any mammal, including a human being.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A-B shows the nucleotide sequence (SEQ ID NO: 1) which comprises the open reading frame encoding the human nuclear receptor protein, nNR5.

Figure 2A-B shows the coding strand of the isolated cDNA molecule (SEQ ID NO: 1) which encodes nNR5, and the amino acid sequence (SEQ ID NO: 2) of nNR5. The region in bold is the DNA binding domain.

Figure 3 shows the amino acid sequence (SEQ ID NO: 2) of nNR5. The region in bold is the DNA binding domain.

DETAILED DESCRIPTION OF THE INVENTION

The present invention relates to isolated nucleic acid and protein forms which represent nuclear receptors, preferably but not necessarily limited to human receptors. These expressed proteins are novel nuclear receptors and which are useful in the identification of downstream target genes and ligands regulating their activity. The nuclear receptor proteins encoded by the isolated polynucleotides of the present invention are involved in the regulation of *in vivo* cell proliferation and/or cell development. The nuclear receptor superfamily is composed of a group of structurally related receptors which are regulated by chemically distinct ligands. The common structure for a nuclear receptor is a highly conserved DNA binding domain (DBD) located in the center of the peptide and the ligand-binding domain (LBD) at the COOH-terminus. Eight out of the nine non-variant cysteines form two type II zinc fingers which distinguish nuclear receptors from other DNA-binding proteins. The DBDs share at least 50% to 60% amino acid sequence identity even among the most distant members in vertebrates. The superfamily has been expanded within the past decade to contain approximately 25 subfamilies. An EST database search using whole peptide sequences of several representative subfamily members, were utilized to identify a human EST (GenBank Acc. No. W27871; dbEST

Id 534939; search available through National Center for Biotechnology Information - http://www.ncbi.nlm.nih.gov/dbEST/index.html) which encodes a portion of a novel member of the nuclear receptor superfamily. In addition, the exemplified cDNA encoding nNR5 was isolated using DNA fragments encoding DBD regions of androgen receptor (AR), estrogen receptor b (ERb), glucocorticoid receptor (GR) and vitamin D receptor (VDR) as probes to screen a human retina cDNA library and a library made from mRNA derived from 20 major human tissues commercially available from Clontech (Palo Alto, CA) at low stringency. Twenty positive clones were obtained by screening 250,000 primary clones from a human retina cDNA library constructed in the lab. Sequence information was obtained by directly sequencing one of the purified clones (Figure 1A-B; SEQ ID NO: 1). A peptide of 367 amino acids encoded by the cDNA has the authentic domain structures of the nuclear receptor (Figure 2A-B, Figure 3; SEQ ID NO: 2). A data base search revealed that two other ESTs from a retina library matching this clone in non-conserved region, which are Gen Bank Acc. No. W21793 (dbEST Id 534939; http://www.ncbi.nlm.nih.gov/dbEST/index.html) and Gen Bank Acc. No. W21801 (dbEST Id 534939; http://www.ncbi.nlm.nih.gov/dbEST/index.html). A known gene which is most related to nNR5 at peptide sequence level is chicken ovalbumin upstream promoter transcription factor (COUP-TF). The protein nNR5 is 43% homologous in overlapping regions to COUP-TF. The gene encoding human nNR5 is located on chromosome 15. Expression of human nNR5 was not detected in the majority of the tissues examined via RT-PCR, but it is very abundant in retina based on screening results. Therefore, nNR5 represents a new subfamily of the nuclear receptor superfamily because its low homology to other members in the superfamily.

The present invention also relates to isolated nucleic acid fragments of nNR5 (SEQ ID NO: 1) which encode mRNA expressing a biologically active novel human nuclear receptor. Any such nucleic acid fragment will encode either a protein or protein fragment comprising at least an intracellular DNA-binding domain and/or ligand binding domain, domains conserved throughout the human nuclear receptor family domain which exist in nNR5 (SEQ ID NO:2). Any such

polynucleotide includes but is not necessarily limited to nucleotide substitutions, deletions, additions, amino-terminal truncations and carboxy-terminal truncations such that these mutations encode mRNA which express a protein or protein fragment of diagnostic, therapeutic or prophylactic use and would be useful for screening for agonists and/or antagonists for nNR5 function.

The isolated nucleic acid molecule of the present invention may include a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which may be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention may also include a ribonucleic acid molecule (RNA).

The present invention also relates to recombinant vectors and recombinant hosts, both prokaryotic and eukaryotic, which contain the substantially purified nucleic acid molecules disclosed throughout this specification.

A preferred aspect of the present invention is disclosed in Figure 1A-B and SEQ ID NO: 1, a human cDNA encoding a novel nuclear trans-acting receptor protein, nNR5, disclosed as follows:

ATTCGGGACC TTGGGGCAGC TCCTGAGTTC AGACAGAGTT CAGGAAGGGA
GACAGGGGCA CAGAGAGACA GAGGTTCATG GACTGAGGCA AAGGCTGGGC
CAGGCTCAGC AACCCAGGCC TCCCGCAGGC AGGCAGAGGC TGCCCTGTAA
CCCATGGAGA CCAGACCAAC AGCTCTGATG AGCTCCACAG TGGCTGCAGC
TGCGCCTGCA GCTGGGGCTG CCTCCAGGAA GGAGTCTCCA GGCAGATGGG
GCCTGGGGGA GGATCCCACA GGCGTGAGCC CCTCGCTCCA GTGCCGCGTG
TGCGGAGACA GCAGCAGCGG GAAGCACTAT GGCATCTATG CCTGCAACGG
CTGCAGCGGC TTCTTCAAGA GGAGCGTACG GCGGAGGCTC ATCTACAGGT
GCCAGGTGGG GGCAGGGATG TGCCCCGTGG ACAAGGCCCA CCGCAACCAG
TGCCAGGCCT GCCGGCTGAA GAAGTGCCTG CAGGCGGGGA TGAACCAGGA
CGCCGTGCAG AACGAGCGCC AGCCGCGAAG CACAGCCCAG GTCCACCTGG
ACAGCATGGA GTCCAACACT GAGTCCCGGC CGGAGTCCCT GGTGGCTCCC
CCGGCCCCGG CAGGGCGCAG CCCACGGGGC CCCACACCCA TGTCTGCAGC
CAGAGCCCTG GGCCACCACT TCATGGCCAG CCTTATAACA GCTGAAACCT
GTGCTAAGCT GGAGCCAGAG GATGCTGATG AGAATATTGA TGTCACCAGC

AATGACCCTG AGTTCCCCTC CTCTCCATAC TCCTCTTCCT CCCCCTGCGG
CCTGGACAGC ATCCATGAGA CCTCGGCTCG CCTACTCTTC ATGGCCGTCA
AGTGGGCCAA GAACCTGCCT GTGTTCTCCA GCCTGCCCTT CCGGGATCAG
GTGATCCTGC TGGAAGAGGC GTGGAGTGAA CTCTTTCTCC TCGGGGCCAT
CCAGTGGTCT CTGCCTCTGG ACAGCTGTCC TCTGCTGGCA CCGCCCGAGG
CTTCTGCTGC CGGTGGTGCC CAGGGCCGGC TCACGCTGGC CAGCATGGAG
ACGCGTGTCC TGCAGGAAAC TATCTCTCGG TTCCGGGCAT TGGCGGTGGA
CCCCACGGAG TTTGCCTGCA TGAAGGCCTT GGTCCTCTTC AAGCCAGAGA
CGCGGGGCCT GAAGGATCCT GAGCACGTAG AGGCCTTGCA GGACCAGTCC
CAAGTGATGC TGAGCCAGCA CAGCAAGGCC CACCACCCCA GCCAGCCCGT
GAGGTGACCT GAGCATGCGC CCACCCACTC ATCTGTCCCT GACCTCTAAC
CTTTCTCTGC CTCTCCCACA CTCTCCCAGA GCTCACTGAT TAGACAGCAC
AAGGGTCTCA GTTCAACAGC ATACAGCCAA CATCTATGGT GTCCCAGGCA
CAGTGCCAGG CCCCGGGAGT GGGGACCAAG ATGTACATAA GACAAAGCTA
CTGCCTTCTA GAGACAACCG GCAGTGACCT CACTGAAGAC AAAAACTGCC
CTAGCCAGGT ACTGAGGGTT GCATGAATCT GCAGGAGACA GAGATCCCCT
TGCATGGGAA ACATAAAGCA GAATTGGGAG GGACTTTGTG GAGACAGGGC
TGGACTTGAA AGGAAGAAGA AGTCTAAAAG AAAACATCAT TTGCAAAGGG
AGAGAGGGGC AAGCATGATA TGTTGTTAGA ACAGGAGCCC ACTTTGAAGG
TATAACAGGT TCCTGCCAGT GAGAAATGGG GAGAATAAGC CAGAAAAGTA
CCCTAGGACC AGCCCGTTCA GGACTTTGAA TGCCAGCCAA AGGCCACGTC
TGACTTGGGA GGCAGAGGGC AGCTACTGCA GGTTTCCGAG CAGAGGGTCA
TACACAGGGC TGGACCTCAC GCAGACTGGC ATGGCCATGG GTCCAGAGGA
TACTACTGGG AAGGGGATGG CAGCTACTGC CACCTTCCAG ATGGTTCCAT
GGAGTTCTGA TCTTTGGGCA TGGCCAGGGG AAGCAGAAGG GAGACTCTAG
GAGTTGAAAT GGGTCAGACC CGGTGTTTGG GTGAAGGTAA GGAATGAGGG
AAGAGGAGCT CTTTG (SEQ ID NO: 1).

The present invention also relates to a substantially purified form of the novel nuclear trans-acting receptor protein, nNR5, which is shown in Figures 2A-B and Figure 3 and as set forth in SEQ ID NO:2, disclosed as follows:

METRPTALMS STVAAAAPAA GAASRKESPG RWGLGEDPTG VSPSLQCRVC
GDSSSGKHYG IYACNGCSGF FKRSVRRRLI YRCQVGAGMC PVDKAHRNQC
QACRLKKCLQ AGMNQDAVQN ERQPRSTAQV HLDSMESNTE SRPESLVAPP
APAGRSPRGP TPMSAARALG HHFMASLITA ETCAKLEPED ADENIDVTSN

DPEFPSSPYS SSSPCGLDSI HETSARLLFM AVKWAKNLPV FSSLPFRDQV
ILLEEAWSEL FLLGAIQWSL PLDSCPLLAP PEASAAGGAQ GRLTLASMET
RVLQETISRF RALAVDPTEF ACMKALVLFK PETRGLKDPE HVEALQDQSQ
VMLSQHSKAH HPSQPVR (SEQ ID NO:2).

The present invention also relates to biologically functional derivatives and/or mutants of nNR5 as set forth as SEQ ID NO:2, including but not necessarily limited to amino acid substitutions, deletions, additions, amino terminal truncations and carboxy-terminal truncations such that these mutations provide for proteins or protein fragments of diagnostic, therapeutic or prophylactic use and would be useful for screening for agonists and/or antagonists of nNR5 function.

The present invention also relates to an isolated cDNA molecule which comprises the nucleotide sequence which encodes the entire reading frame of human NR5, as well as containing an intron, from nucleotide 971 to nucleotide 1847, as underlined below and as set forth as SEQ ID NO: 18.

TATAGGGCGA ATTGGGTACC GGGCCCCCCC TCGAGGTCGA CGGTATCGAT
AAGCTTGATA TCGAATTCGA ATTCGGGACC TTGGGGCAGC TCCTGAGTTC
AGACAGAGTT CAGGAAGGGA GACAGGGGCA CAGAGAGACA GAGGTTCATG
GACTGAGGCA AAGGCTGGGC CAGGCTCAGC AACCCAGGCC TCCCGCAGGC
AGGCAGAGGC TGCCCTGTAA CCCATGGAGA CCAGACCAAC AGCTCTGATG
AGCTCCACAG TGGCTGCAGC TGCGCCTGCA GCTGGGGCTG CCTCCAGGAA
GGAGTCTCCA GGCAGATGGG GCCTGGGGGA GGATCCCACA GGCGTGAGCC
CCTCGCTCCA GTGCCGCGTG TGCGGAGACA GCAGCAGCGG GAAGCACTAT
GGCATCTATG CCTGCAACGG CTGCAGCGGC TTCTTCAAGA GGAGCGTACG
GCGGAGGCTC ATCTACAGGT GCCAGGTGGG GGCAGGGATG TGCCCCGTGG
ACAAGGCCCA CCGCAACCAG TGCCAGGCCT GCCGGCTGAA GAAGTGCCTG
CAGGCGGGGA TGAACCAGGA CGCCGTGCAG AACGAGCGCC AGCCGCGAAG
CACAGCCCAG GTCCACCTGG ACAGCATGGA GTCCAACACT GAGTCCCGGC
CGGAGTCCCT GGTGGCTCCC CCGGCCCCGG CAGGGCGCAG CCCACGGGGC
CCCACACCCA TGTCTGCAGC CAGAGCCCTG GGCCACCACT TCATGGCCAG
CCTTATAACA GCTGAAACCT GTGCTAAGCT GGAGCCAGAG GATGCTGATG
AGAATATTGA TGTCACCAGC AATGACCCTG AGTTCCCCTC CTCTCCATAC
TCCTCTTCCT CCCCCTGCGG CCTGGACAGC ATCCATGAGA CCTCGGCTCG
CCTACTCTTC ATGGCCGTCA AGTGGGCCAA GAACCTGCCT GTGTTCTCCA

GCCTGCCCTT CCGGGATCAG GTACCTACCG GCCTGCCTGC TGGGGAGCTA
GGCTGGGCTG GGGTCAGGCG GCCCACTCGA GTCAACCAGA CAGGGCACAC
ACATCCCCAC GCCAGTATGA ATGCACACAG CTTGGATGGT GATGGCTGGG
GACACACATA CCTCTGATTC AGCGATGGCT GGGGTGCATC TCAGGGATGG
TGACGGTGGG GGTGCATGCA TCTCTGGCAC AGGGATGATG GTCGGGGTGC
ACACCTAGGA GATGATGATG GCTAGGGACC TACAGGGCCC AGGGTCTTCT
TAAGTTCTGG AAGACCCTCA GGCCCTGCAG ACATTCTGTG GGTAACAAGT
GACCTGCACA CCCTGAACAG GCTGAGTGGC TGACTCTAGG CCCCCTTGGA
GCACAAGTGC CTACGACTTC AGGGCTTGCA TTTTAGTTCA ATCTCTCCAG
CTCTGGGCCA TCCCTCTCGG CTTCTAATGG GCAAGCAGAT CTTTCAGGAA
AACCAGGAGG AGAGGCATGA GGAAGGTTTG AGGCCCTCAG CCAGTCTGTG
TGCTGGGGTG GAGCAACTCA GAAGAGTCAG GCCACACCAC TTGAATACAC
TCAACTTAGG ACACTCATGA GGCATGTCTC TGAGGCTGCC CAACTTCCAA
TGGCTCTGGG CGTTCCTAAA TGTCCCAGCT GCAGCTCTGG ATGGAACCCA
GTGTCTCAGA TGATAGGCAG CTGAGCCGGA TGGTGCCAAA TCCCAGAGCT
CTGAGCCTCT GGCTGATGTC AGGAGAGCAT TCTCGGGTCC CAGGACAGCA
CTTCCATTCC TTGGGTGCCT GAGATGGTGG CAGAGGCTCC AGACTGAGCC
AGAGAAGCTG TGTGTCTGCC ATAACAGGCA CCCCTGTCTG AGCACAGGTG
ATCCTGCTGG AAGAGGCGTG GAGTGAACTC TTTCTCCTCG GGGCCATCCA
GTGGTCTCTG CCTCTGGACA GCTGTCCTCT GCTGGCACCG CCCGAGGCCT
CTGCTGCCGG TGGTGCCCAG GGCCGGCTCA CGCTGGCCAG CATGGAGACG
CGTGTCCTGC AGGAAACTAT CTCTCGGTTC CGGGCATTGG CGGTGGACCC
CACGGAGTTT GCCTGCATGA AGGCCTTGGT CCTCTTCAAG CCAGAGACGC
GGGGCCTGAA GGATCCTGAG CACGTAGAGG CCTTGCAGGA CCAGTCCCAA
GTGATGCTGA GCCAGCACAG CAAGGCCCAC CACCCCAGCC AGCCCGTGAG
GTGACCTGAG CATGCGCCCA CCCACTCATC TGTCCCTGAC CTCTAACCTT
TCTCTGCCTC TCCCACACTC TCCCAGAGCT CACTGATTAG ACAGCACAAG
GGTCTCAGTT CAACAGCATA CAGCCAACAT CTATGGTGTC CCAGGCACAG
TGCCAGGCCC CGGGAGTGGG GACCAAGATG TACATAAGAC AAAGCTACTG
CCTTCTAGAG ACAACCGGCA GTGACCTCAC TGAAGACAAA AACTGCCCTA
GCCAGGTACT GAGGGTTGCA TGAATCTGCA GGAGACAGAG ATCCCCTTGC
ATGGGAAACA TAAAGCAGAA TTGGGAGGGA CTTTGTGGAG ACAGGGCTGG
ACTTGAAAGG AAGAAGAAGT CTAAAAGAAA ACATCATTTG CAAAGGGAGA
GAGGGGCAAG CATGATATGT TGTTAGAACA GGAGCCCACT TTGAAGGTAT
AACAGGTTCC TGCCAGTGAG AAATGGGGAG AATAAGCCAG AAAAGTACCC

TAGGACCAGC CCGTTCAGGA CTTTGAATGC CAGCCAAAGG CCACGTCTGA
CTTGGGAGGC AGAGGGCAGC TACTGCAGGT TTCCGAGCAG AGGGTCATAC
ACAGGGCTGG ACCTCACGCA GACTGGCATG GCCATGGGTC CAGAGGATAC
TACTGGGAAG GGGATGGCAG CTACTGCCAC CTTCCAGATG GTTCCATGGA
GTTCTGATCT TTGGGCATGG CCAGGGGAAG CAGAAGGGAG ACTCTAGGAG
TTGAAATGGG TCAGACCCGG TGTTTGGGTG AAGGTAAGGA ATGAGGGAAG
AGGAGCTCTT TG (SEQ ID NO: 18).

The intron-containing nNR5 cDNA as set forth in SEQ ID NO: 18 contains an additional 70 nucleotides at the 5' end of the clone. Therefore, the present invention also relates to an isolated cDNA which comprises the open reading frame of SEQ ID NO:1, in addition to the additional 70 nucleotides at the 5' end of an isolated polynucleotide encoding nNR5. This nucleotide sequence is shown below and is as set forth in SEQ ID NO: 19:

TATAGGGCGA ATTGGGTACC GGGCCCCCCC TCGAGGTCGA CGGTATCGAT
AAGCTTGATA TCGAATTCGA ATTCGGGACC TTGGGGCAGC TCCTGAGTTC
AGACAGAGTT CAGGAAGGGA GACAGGGGCA CAGAGAGACA GAGGTTCATG
GACTGAGGCA AAGGCTGGGC CAGGCTCAGC AACCCAGGCC TCCCGCAGGC
AGGCAGAGGC TGCCCTGTAA CCCATGGAGA CCAGACCAAC AGCTCTGATG
AGCTCCACAG TGGCTGCAGC TGCGCCTGCA GCTGGGGCTG CCTCCAGGAA
GGAGTCTCCA GGCAGATGGG GCCTGGGGGA GGATCCCACA GGCGTGAGCC
CCTCGCTCCA GTGCCGCGTG TGCGGAGACA GCAGCAGCGG GAAGCACTAT
GGCATCTATG CCTGCAACGG CTGCAGCGGC TTCTTCAAGA GGAGCGTACG
GCGGAGGCTC ATCTACAGGT GCCAGGTGGG GGCAGGGATG TGCCCCGTGG
ACAAGGCCCA CCGCAACCAG TGCCAGGCCT GCCGGCTGAA GAAGTGCCTG
CAGGCGGGGA TGAACCAGGA CGCCGTGCAG AACGAGCGCC AGCCGCGAAG
CACAGCCCAG GTCCACCTGG ACAGCATGGA GTCCAACACT GAGTCCCGGC
CGGAGTCCCT GGTGGCTCCC CCGGCCCCGG CAGGGCGCAG CCCACGGGGC
CCCACACCCA TGTCTGCAGC CAGAGCCCTG GGCCACCACT TCATGGCCAG
CCTTATAACA GCTGAAACCT GTGCTAAGCT GGAGCCAGAG GATGCTGATG
AGAATATTGA TGTCACCAGC AATGACCCTG AGTTCCCCTC CTCTCCATAC
TCCTCTTCCT CCCCCTGCGG CCTGGACAGC ATCCATGAGA CCTCGGCTCG
CCTACTCTTC ATGGCCGTCA AGTGGGCCAA GAACCTGCCT GTGTTCTCCA
GCCTGCCCTT CCGGGATCAG GTGATCCTGC TGGAAGAGGC GTGGAGTGAA
CTCTTTCTCC TCGGGGCCAT CCAGTGGTCT CTGCCTCTGG ACAGCTGTCC

TCTGCTGGCA CCGCCCGAGG CCTCTGCTGC CGGTGGTGCC CAGGGCCGGC
TCACGCTGGC CAGCATGGAG ACGCGTGTCC TGCAGGAAAC TATCTCTCGG
TTCCGGGCAT TGGCGGTGGA CCCCACGGAG TTTGCCTGCA TGAAGGCCTT
GGTCCTCTTC AAGCCAGAGA CGCGGGGCCT GAAGGATCCT GAGCACGTAG
AGGCCTTGCA GGACCAGTCC CAAGTGATGC TGAGCCAGCA CAGCAAGGCC
CACCACCCCA GCCAGCCCGT GAGGTGACCT GAGCATGCGC CCACCCACTC
ATCTGTCCCT GACCTCTAAC CTTTCTCTGC CTCTCCCACA CTCTCCCAGA
GCTCACTGAT TAGACAGCAC AAGGGTCTCA GTTCAACAGC ATACAGCCAA
CATCTATGGT GTCCCAGGCA CAGTGCCAGG CCCCGGGAGT GGGGACCAAG
ATGTACATAA GACAAAGCTA CTGCCTTCTA GAGACAACCG GCAGTGACCT
CACTGAAGAC AAAAACTGCC CTAGCCAGGT ACTGAGGGTT GCATGAATCT
GCAGGAGACA GAGATCCCCT TGCATGGGAA ACATAAAGCA GAATTGGGAG
GGACTTTGTG GAGACAGGGC TGGACTTGAA AGGAAGAAGA AGTCTAAAAG
AAAACATCAT TTGCAAAGGG AGAGAGGGGC AAGCATGATA TGTTGTTAGA
ACAGGAGCCC ACTTTGAAGG TATAACAGGT TCCTGCCAGT GAGAAATGGG
GAGAATAAGC CAGAAAAGTA CCCTAGGACC AGCCCGTTCA GGACTTTGAA
TGCCAGCCAA AGGCCACGTC TGACTTGGGA GGCAGAGGGC AGCTACTGCA
GGTTTCCGAG CAGAGGGTCA TACACAGGGC TGGACCTCAC GCAGACTGGC
ATGGCCATGG GTCCAGAGGA TACTACTGGG AAGGGGATGG CAGCTACTGC
CACCTTCCAG ATGGTTCCAT GGAGTTCTGA TCTTTGGGCA TGGCCAGGGG
AAGCAGAAGG GAGACTCTAG GAGTTGAAAT GGGTCAGACC CGGTGTTTGG
GTGAAGGTAA GGAATGAGGG AAGAGGAGCT CTTTG (SEQ ID NO: 19).

The present invention also relates to isolated nucleic acid molecules which are fusion constructions expressing fusion proteins useful in assays to identify compounds which modulate wild-type human nNR5 activity. A preferred aspect of this portion of the invention includes, but is not limited to, glutathione S-transferase GST-nNR5 fusion constructs. These fusion constructs include, but are not limited to, all or a portion of the ligand-binding domain of nNR5, respectively, as an in-frame fusion at the carboxy terminus of the GST gene. The disclosure of SEQ ID NOS:1-2 allow the artisan of ordinary skill to construct any such nucleic acid molecule encoding a GST-nuclear receptor fusion protein. Soluble recombinant GST-nuclear receptor fusion proteins may be expressed in various expression systems,

including *Spodoptera frugiperda* (Sf21) insect cells (Invitrogen) using a baculovirus expression vector (e.g., Bac-N-Blue DNA from Invitrogen or pAcG2T from Pharmingen).

The isolated nucleic acid molecule of the present invention may include a deoxyribonucleic acid molecule (DNA), such as genomic DNA and complementary DNA (cDNA), which may be single (coding or noncoding strand) or double stranded, as well as synthetic DNA, such as a synthesized, single stranded polynucleotide. The isolated nucleic acid molecule of the present invention may also include a ribonucleic acid molecule (RNA).

It is known that there is a substantial amount of redundancy in the various codons which code for specific amino acids. Therefore, this invention is also directed to those DNA sequences encode RNA comprising alternative codons which code for the eventual translation of the identical amino acid, as shown below:

A=Ala=Alanine: codons GCA, GCC, GCG, GCU
C=Cys=Cysteine: codons UGC, UGU
D=Asp=Aspartic acid: codons GAC, GAU
E=Glu=Glutamic acid: codons GAA, GAG
F=Phe=Phenylalanine: codons UUC, UUU
G=Gly=Glycine: codons GGA, GGC, GGG, GGU
H=His =Histidine: codons CAC, CAU
I=Ile =Isoleucine: codons AUA, AUC, AUU
K=Lys=Lysine: codons AAA, AAG
L=Leu=Leucine: codons UUA, UUG, CUA, CUC, CUG, CUU
M=Met=Methionine: codon AUG
N=Asp=Asparagine: codons AAC, AAU
P=Pro=Proline: codons CCA, CCC, CCG, CCU
Q=Gln=Glutamine: codons CAA, CAG
R=Arg=Arginine: codons AGA, AGG, CGA, CGC, CGG, CGU
S=Ser=Serine: codons AGC, AGU, UCA, UCC, UCG, UCU
T=Thr=Threonine: codons ACA, ACC, ACG, ACU
V=Val=Valine: codons GUA, GUC, GUG, GUU
W=Trp=Tryptophan: codon UGG
Y=Tyr=Tyrosine: codons UAC, UAU.

Therefore, the present invention discloses codon redundancy which may result in differing DNA molecules expressing an identical protein. For purposes of this specification, a sequence bearing one or more replaced codons will be defined as a degenerate variation. Also included within the scope of this invention are mutations either in the DNA sequence or the translated protein which do not substantially alter the ultimate physical properties of the expressed protein. For example, substitution of valine for leucine, arginine for lysine, or asparagine for glutamine may not cause a change in functionality of the polypeptide.

It is known that DNA sequences coding for a peptide may be altered so as to code for a peptide having properties that are different than those of the naturally occurring peptide. Methods of altering the DNA sequences include but are not limited to site directed mutagenesis. Examples of altered properties include but are not limited to changes in the affinity of an enzyme for a substrate or a receptor for a ligand.

As used herein, "purified" and "isolated" are utilized interchangeably to stand for the proposition that the nucleic acid, protein, or respective fragment thereof in question has been substantially removed from its *in vivo* environment so that it may be manipulated by the skilled artisan, such as but not limited to nucleotide sequencing, restriction digestion, site-directed mutagenesis, and subcloning into expression vectors for a nucleic acid fragment as well as obtaining the protein or protein fragment in pure quantities so as to afford the opportunity to generate polyclonal antibodies, monoclonal antibodies, amino acid sequencing, and peptide digestion. Therefore, the nucleic acids claimed herein may be present in whole cells or in cell lysates or in a partially purified or substantially purified form. A nucleic acid is considered substantially purified when it is purified away from environmental contaminants. Thus, a nucleic acid sequence isolated from cells is considered to be substantially purified when purified from cellular components by standard methods while a

chemically synthesized nucleic acid sequence is considered to be substantially purified when purified from its chemical precursors.

The present invention also relates to recombinant vectors and recombinant hosts, both prokaryotic and eukaryotic, which contain the substantially purified nucleic acid molecules disclosed throughout this specification.

Therefore, the present invention also relates to methods of expressing nNR5 and biological equivalents disclosed herein, assays employing these recombinantly expressed gene products, cells expressing these gene products, and agonistic and/or antagonistic compounds identified through the use of assays utilizing these recombinant forms, including, but not limited to, one or more modulators of the human nNR5 either through direct contact LBD or through direct or indirect contact with a ligand which either interacts with the DBD or with the wild-type transcription complex which nNR5 interacts *in trans*, thereby modulating cell differentiation or cell development.

As used herein, a "biologically functional derivative" of a wild-type human nNR5 possesses a biological activity that is related to the biological activity of the wild type human nNR5 . The term "functional derivative" is intended to include the "fragments," "mutants," "variants," "degenerate variants," "analogs" and "homologues" of the wild type human nNR5 protein. The term "fragment" is meant to refer to any polypeptide subset of wild-type human nNR5, including but not necessarily limited to nNR5 proteins comprising amino acid substitutions, deletions, additions, amino terminal truncations and/or carboxy-terminal truncations. The term "mutant" is meant to refer a subset of a biologically active fragment that may be substantially similar to the wild-type form but possesses distinguishing biological characteristics. Such altered characteristics include but are in no way limited to altered substrate binding, altered substrate affinity and altered sensitivity to chemical compounds affecting biological activity of the human nNR5 or human nNR5 functional derivative. The term "variant" is meant to refer to a molecule substantially similar in structure and function to either the entire wild-

type protein or to a fragment thereof. A molecule is "substantially similar" to a wild-type human nNR5-like protein if both molecules have substantially similar structures or if both molecules possess similar biological activity. Therefore, if the two molecules possess substantially similar activity, they are considered to be variants even if the structure of one of the molecules is not found in the other or even if the two amino acid sequences are not identical. The term "analog" refers to a molecule substantially similar in function to either the full-length human nNR5 protein or to a biologically functional derivative thereof.

Any of a variety of procedures may be used to clone human nNR5. These methods include, but are not limited to, (1) a RACE PCR cloning technique (Frohman, et al., 1988, *Proc. Natl. Acad. Sci. USA* 85: 8998-9002). 5′ and/or 3′ RACE may be performed to generate a full length cDNA sequence. This strategy involves using gene-specific oligonucleotide primers for PCR amplification of human nNR5 cDNA. These gene-specific primers are designed through identification of an expressed sequence tag (EST) nucleotide sequence which has been identified by searching any number of publicly available nucleic acid and protein databases; (2) direct functional expression of the human nNR5 cDNA following the construction of a human nNR5-containing cDNA library in an appropriate expression vector system; (3) screening a human nNR5-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a labeled degenerate oligonucleotide probe designed from the amino acid sequence of the human nNR5 protein; (4) screening a human nNR5-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a partial cDNA encoding the human nNR5 protein. This partial cDNA is obtained by the specific PCR amplification of human nNR5 DNA fragments through the design of degenerate oligonucleotide primers from the amino acid sequence known for other kinases which are related to the human nNR5 protein; (5) screening a human nNR5-containing cDNA library constructed in a bacteriophage or plasmid shuttle vector with a partial cDNA encoding the human nNR5 protein. This strategy may also involve using gene-specific oligonucleotide primers for PCR amplification of human nNR5 cDNA identified as an EST as described above; or (6) designing 5′ and 3′

gene specific oligonucleotides using SEQ ID NO: 1 as a template so that either the full-length cDNA may be generated by known PCR techniques, or a portion of the coding region may be generated by these same known PCR techniques to generate and isolate a portion of the coding region to use as a probe to screen one of numerous types of cDNA and/or genomic libraries in order to isolate a full-length version of the nucleotide molecule encoding human nNR5 .

It is readily apparent to those skilled in the art that other types of libraries, as well as libraries constructed from other cell types-or species types, may be useful for isolating a nNR5-encoding DNA or a nNR5 homologue. Other types of libraries include, but are not limited to, cDNA libraries derived from other cells or cell lines other than human cells or tissue such as murine cells, rodent cells or any other such vertebrate host which may contain nNR5-encoding DNA. Additionally a nNR5 gene and homologues may be isolated by oligonucleotide- or polynucleotide-based hybridization screening of a vertebrate genomic library, including but not limited to, a murine genomic library, a rodent genomic library, as well as concomitant human genomic DNA libraries.

It is readily apparent to those skilled in the art that suitable cDNA libraries may be prepared from cells or cell lines which have nNR5 activity. The selection of cells or cell lines for use in preparing a cDNA library to isolate a cDNA encoding nNR5 may be done by first measuring cell-associated nNR5 activity using any known assay available for such a purpose.

Preparation of cDNA libraries can be performed by standard techniques well known in the art. Well known cDNA library construction techniques can be found for example, in Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory, Cold Spring Harbor, New York. Complementary DNA libraries may also be obtained from numerous commercial sources, including but not limited to Clontech Laboratories, Inc. and Stratagene.

It is also readily apparent to those skilled in the art that DNA encoding human nNR5 may also be isolated from a suitable genomic DNA library. Construction of genomic DNA libraries can be

performed by standard techniques well known in the art. Well known genomic DNA library construction techniques can be found in Sambrook, et al., *supra*.

In order to clone the human nNR5 gene by one of the preferred methods, the amino acid sequence or DNA sequence of human nNR5 or a homologous protein may be necessary. To accomplish this, the nNR5 protein or a homologous protein may be purified and partial amino acid sequence determined by automated sequenators. It is not necessary to determine the entire amino acid sequence, but the linear sequence of two regions of 6 to 8 amino acids can be determined for the PCR amplification of a partial human nNR5 DNA fragment. Once suitable amino acid sequences have been identified, the DNA molecules capable of encoding them are synthesized. Because the genetic code is degenerate, more than one codon may be used to encode a particular amino acid, and therefore, the amino acid sequence can be encoded by any of a set of similar DNA oligonucleotides. Only one member of the set will be identical to the human nNR5 sequence but others in the set will be capable of hybridizing to human nNR5 DNA even in the presence of DNA oligonucleotides with mismatches. The mismatched DNA oligonucleotides may still sufficiently hybridize to the human nNR5 DNA to permit identification and isolation of human nNR5 encoding DNA. Alternatively, the nucleotide sequence of a region of an expressed sequence may be identified by searching one or more available genomic databases. Gene-specific primers may be used to perform PCR amplification of a cDNA of interest from either a cDNA library or a population of cDNAs. As noted above, the appropriate nucleotide sequence for use in a PCR-based method may be obtained from SEQ ID NO: 1, either for the purpose of isolating overlapping 5′ and 3′ RACE products for generation of a full-length sequence coding for human nNR5, or to isolate a portion of the nucleotide molecule coding for human nNR5 for use as a probe to screen one or more cDNA- or genomic-based libraries to isolate a full-length molecule encoding human nNR5 or human nNR5-like proteins.

In an exemplified method, the human nNR5 full-length cDNA of the present invention was isolated by screening a human retina cDNA library with an oligonucleotide primer pair to a human EST identified herein as SEQ ID NO: 3. Positive cDNA clones were sequenced and shown to possess an intron. This cDNA was subjected to sequence analysis and is reported herein and is set forth as SEQ ID NO: 18. A second oligonucleotide primer pair which flanks the putative intron was used to rescreen the human retina cDNA library. Shorter cDNA clones (about 2.1 kb) were chosen for sequence analysis and shown to comprise an uninterrupted open reading frame (e.g., SEQ ID NO:1) encoding human nNR5 (SEQ ID NO: 2). The intron-containing clone disclosed as SEQ ID NO: 18 contains 70 additional nucleotides at the 5' end of the cDNA clone. Therefore, an additional isolated DNA molecule of the present invention includes but is not limited to the DNA molecule as set forth herein and as set forth as SEQ ID NO: 19.

A variety of mammalian expression vectors may be used to express recombinant human nNR5 in mammalian cells. Expression vectors are defined herein as DNA sequences that are required for the transcription of cloned DNA and the translation of their mRNAs in an appropriate host. Such vectors can be used to express eukaryotic DNA in a variety of hosts such as bacteria, blue green algae, plant cells, insect cells and animal cells. Specifically designed vectors allow the shuttling of DNA between hosts such as bacteria-yeast or bacteria-animal cells. An appropriately constructed expression vector should contain: an origin of replication for autonomous replication in host cells, selectable markers, a limited number of useful restriction enzyme sites, a potential for high copy number, and active promoters. A promoter is defined as a DNA sequence that directs RNA polymerase to bind to DNA and initiate RNA synthesis. A strong promoter is one which causes mRNAs to be initiated at high frequency. Expression vectors may include, but are not limited to, cloning vectors, modified cloning vectors, specifically designed plasmids or viruses.

Commercially available mammalian expression vectors which may be suitable for recombinant human nNR5 expression, include but are not limited to, pcDNA3.1 (Invitrogen), pLITMUS28,

pLITMUS29, pLITMUS38 and pLITMUS39 (New England Bioloabs), pcDNAI, pcDNAIamp (Invitrogen), pcDNA3 (Invitrogen), pMC1neo (Stratagene), pXT1 (Stratagene), pSG5 (Stratagene), EBO-pSV2-neo (ATCC 37593) pBPV-1(8-2) (ATCC 37110), pdBPV-MMTneo(342-12) (ATCC 37224), pRSVgpt (ATCC 37199), pRSVneo (ATCC 37198), pSV2-dhfr (ATCC 37146), pUCTag (ATCC 37460), and lZD35 (ATCC 37565).

A variety of bacterial expression vectors may be used to express recombinant human nNR5 in bacterial cells. Commercially available bacterial expression vectors which may be suitable for recombinant human nNR5 expression include, but are not limited to pCRII (Invitrogen), pCR2.1 (Invitrogen), pQE (Qiagen), pET11a (Novagen), lambda gt11 (Invitrogen), and pKK223-3 (Pharmacia).

A variety of fungal cell expression vectors may be used to express recombinant human nNR5 in fungal cells. Commercially available fungal cell expression vectors which may be suitable for recombinant human nNR5 expression include but are not limited to pYES2 (Invitrogen) and *Pichia* expression vector (Invitrogen).

A variety of insect cell expression vectors may be used to express recombinant receptor in insect cells. Commercially available insect cell expression vectors which may be suitable for recombinant expression of human nNR5 include but are not limited to pBlueBacIII and pBlueBacHis2 (Invitrogen), and pAcG2T (Pharmingen).

An expression vector containing DNA encoding a human nNR5-like protein may be used for expression of human nNR5 in a recombinant host cell. Recombinant host cells may be prokaryotic or eukaryotic, including but not limited to bacteria such as *E. coli*, fungal cells such as yeast, mammalian cells including but not limited to cell lines of human, bovine, porcine, monkey and rodent origin, and insect cells including but not limited to Drosophila- and silkworm-derived cell lines. Cell lines derived from mammalian species which may be suitable and which are commercially available, include but are not limited to, L cells L-M(TK$^-$) (ATCC CCL 1.3), L cells L-M (ATCC CCL 1.2), Saos-2 (ATCC HTB-85), 293 (ATCC CRL 1573), Raji (ATCC CCL 86), CV-1 (ATCC CCL 70), COS-1 (ATCC CRL 1650), COS-7 (ATCC CRL 1651), CHO-K1 (ATCC CCL 61), 3T3 (ATCC CCL 92), NIH/3T3 (ATCC

CRL 1658), HeLa (ATCC CCL 2), C127I (ATCC CRL 1616), BS-C-1 (ATCC CCL 26), MRC-5 (ATCC CCL 171) and CPAE (ATCC CCL 209).

The expression vector may be introduced into host cells via any one of a number of techniques including but not limited to transformation, transfection, protoplast fusion, and electroporation. The expression vector-containing cells are individually analyzed to determine whether they produce human nNR5 protein. Identification of human nNR5 expressing cells may be done by several means, including but not limited to immunological reactivity with anti-human nNR5 antibodies, labeled ligand binding and the presence of host cell-associated human nNR5 activity.

The cloned human nNR5 cDNA obtained through the methods described above may be recombinantly expressed by molecular cloning into an expression vector (such as pcDNA3.1, pQE, pBlueBacHis2 and pLITMUS28) containing a suitable promoter and other appropriate transcription regulatory elements, and transferred into prokaryotic or eukaryotic host cells to produce recombinant human nNR5. Techniques for such manipulations can be found described in Sambrook, et al., *supra* , are discussed at length in the Example section and are well known and easily available to the artisan of ordinary skill in the art.

Expression of human nNR5 DNA may also be performed using *in vitro* produced synthetic mRNA. Synthetic mRNA can be efficiently translated in various cell-free systems, including but not limited to wheat germ extracts and reticulocyte extracts, as well as efficiently translated in cell based systems, including but not limited to microinjection into frog oocytes, with microinjection into frog oocytes being preferred.

To determine the human nNR5 cDNA sequence(s) that yields optimal levels of human nNR5, cDNA molecules including but not limited to the following can be constructed: a cDNA fragment containing the full-length open reading frame for human nNR5 as well as various constructs containing portions of the cDNA encoding only specific domains of the protein or rearranged domains of the protein. All constructs can be designed to contain none, all or portions of the 5′

and/or 3′ untranslated region of a human nNR5 cDNA. The expression levels and activity of human nNR5 can be determined following the introduction, both singly and in combination, of these constructs into appropriate host cells. Following determination of the human nNR5 cDNA cassette yielding optimal expression in transient assays, this nNR5 cDNA construct is transferred to a variety of expression vectors (including recombinant viruses), including but not limited to those for mammalian cells, plant cells, insect cells, oocytes, bacteria, and yeast cells.

The present invention also relates to polyclonal and monoclonal antibodies raised in response to either the human form of nNR5 disclosed herein, or a biologically functional derivative thereof. It will be especially preferable to raise antibodies against epitopes within the NH_2-terminal domain of nNR5, which show the least homology to other known proteins belonging to the human nuclear receptor superfamily.

Recombinant nNR5 protein can be separated from other cellular proteins by use of an immunoaffinity column made with monoclonal or polyclonal antibodies specific for full-length nNR5 protein, or polypeptide fragments of nNR5 protein. Additionally, polyclonal or monoclonal antibodies may be raised against a synthetic peptide (usually from about 9 to about 25 amino acids in length) from a portion of the protein as disclosed in SEQ ID NO:2. Monospecific antibodies to human nNR5 are purified from mammalian antisera containing antibodies reactive against human nNR5 or are prepared as monoclonal antibodies reactive with human nNR5 using the technique of Kohler and Milstein (1975, *Nature* 256: 495-497). Monospecific antibody as used herein is defined as a single antibody species or multiple antibody species with homogenous binding characteristics for human nNR5. Homogenous binding as used herein refers to the ability of the antibody species to bind to a specific antigen or epitope, such as those associated with human nNR5, as described above. Human nNR5-specific antibodies are raised by immunizing animals such as mice, rats, guinea pigs, rabbits, goats, horses and the like, with an

appropriate concentration of human nNR5 protein or a synthetic peptide generated from a portion of human nNR5 with or without an immune adjuvant.

Preimmune serum is collected prior to the first immunization. Each animal receives between about 0.1 mg and about 1000 mg of human nNR5 protein associated with an acceptable immune adjuvant. Such acceptable adjuvants include, but are not limited to, Freund's complete, Freund's incomplete, alum-precipitate, water in oil emulsion containing *Corynebacterium parvum* and tRNA. The initial immunization consists of human nNR5 protein or peptide fragment thereof in, preferably, Freund's complete adjuvant at multiple sites either subcutaneously (SC), intraperitoneally (IP) or both. Each animal is bled at regular intervals, preferably weekly, to determine antibody titer. The animals may or may not receive booster injections following the initial immunization. Those animals receiving booster injections are generally given an equal amount of human nNR5 in Freund's incomplete adjuvant by the same route. Booster injections are given at about three week intervals until maximal titers are obtained. At about 7 days after each booster immunization or about weekly after a single immunization, the animals are bled, the serum collected, and aliquots are stored at about -20°C.

Monoclonal antibodies (mAb) reactive with human nNR5 are prepared by immunizing inbred mice, preferably Balb/c, with human nNR5 protein. The mice are immunized by the IP or SC route with about 1 mg to about 100 mg, preferably about 10 mg, of human nNR5 protein in about 0.5 ml buffer or saline incorporated in an equal volume of an acceptable adjuvant, as discussed above. Freund's complete adjuvant is preferred. The mice receive an initial immunization on day 0 and are rested for about 3 to about 30 weeks. Immunized mice are given one or more booster immunizations of about 1 to about 100 mg of human nNR5 in a buffer solution such as phosphate buffered saline by the intravenous (IV) route. Lymphocytes, from antibody positive mice, preferably splenic lymphocytes, are obtained by removing spleens from immunized mice by standard procedures known in the art. Hybridoma cells are produced by mixing the splenic

lymphocytes with an appropriate fusion partner, preferably myeloma cells, under conditions which will allow the formation of stable hybridomas. Fusion partners may include, but are not limited to: mouse myelomas P3/NS1/Ag 4-1, MPC-11, S-194 and Sp 2/0, with Sp 2/0 being preferred. The antibody producing cells and myeloma cells are fused in polyethylene glycol, about 1000 mol. wt., at concentrations from about 30% to about 50%. Fused hybridoma cells are selected by growth in hypoxanthine, thymidine and aminopterin supplemented Dulbecco's Modified Eagles Medium (DMEM) by procedures known in the art. Supernatant fluids are collected form growth positive wells on about days 14, 18, and 21 and are screened for antibody production by an immunoassay such as solid phase immunoradioassay (SPIRA) using human nNR5 as the antigen. The culture fluids are also tested in the Ouchterlony precipitation assay to determine the isotype of the mAb. Hybridoma cells from antibody positive wells are cloned by a technique such as the soft agar technique of MacPherson, 1973, Soft Agar Techniques, in *Tissue Culture Methods and Applications*, Kruse and Paterson, Eds., Academic Press.

Monoclonal antibodies are produced *in vivo* by injection of pristine primed Balb/c mice, approximately 0.5 ml per mouse, with about 2 x 10^6 to about 6 x 10^6 hybridoma cells about 4 days after priming. Ascites fluid is collected at approximately 8-12 days after cell transfer and the monoclonal antibodies are purified by techniques known in the art.

In vitro production of anti-human nNR5 mAb is carried out by growing the hybridoma in DMEM containing about 2% fetal calf serum to obtain sufficient quantities of the specific mAb. The mAb are purified by techniques known in the art.

Antibody titers of ascites or hybridoma culture fluids are determined by various serological or immunological assays which include, but are not limited to, precipitation, passive agglutination, enzyme-linked immunosorbent antibody (ELISA) technique and radioimmunoassay (RIA) techniques. Similar assays are used to detect the presence of human nNR5 in body fluids or tissue and cell extracts.

It is readily apparent to those skilled in the art that the above described methods for producing monospecific antibodies may be utilized to produce antibodies specific for human nNR5 peptide fragments, or full-length human nNR5.

Human nNR5 antibody affinity columns are made, for example, by adding the antibodies to Affigel-10 (Biorad), a gel support which is pre-activated with N-hydroxysuccinimide esters such that the antibodies form covalent linkages with the agarose gel bead support. The antibodies are then coupled to the gel via amide bonds with the spacer arm. The remaining activated esters are then quenched with 1M ethanolamine HCl (pH 8.0). The column is washed with water followed by 0.23 M glycine HCl (pH 2.6) to remove any non-conjugated antibody or extraneous protein. The column is then equilibrated in phosphate buffered saline (pH 7.3) and the cell culture supernatants or cell extracts containing full-length human nNR5 or human nNR5 protein fragments are slowly passed through the column. The column is then washed with phosphate buffered saline until the optical density (A280) falls to background, then the protein is eluted with 0.23 M glycine-HCl (pH 2.6). The purified human nNR5 protein is then dialyzed against phosphate buffered saline.

Levels of human nNR5 in host cells is quantified by a variety of techniques including, but not limited to, immunoaffinity and/or ligand affinity techniques. nNR5-specific affinity beads or nNR5-specific antibodies are used to isolate ^{35}S-methionine labeled or unlabelled nNR5. Labeled nNR5 protein is analyzed by SDS-PAGE. Unlabelled nNR5 protein is detected by Western blotting, ELISA or RIA assays employing either nNR5 protein specific antibodies and/or antiphosphotyrosine antibodies.

Following expression of nNR5 in a host cell, nNR5 protein may be recovered to provide nNR5 protein in active form. Several nNR5 protein purification procedures are available and suitable for use. Recombinant nNR5 protein may be purified from cell lysates and extracts, or from conditioned culture medium, by various combinations of, or individual application of salt fractionation, ion exchange chromatography, size exclusion chromatography, hydroxylapatite

adsorption chromatography and hydrophobic interaction chromatography.

The present invention is also directed to methods for screening for compounds which modulate the expression of DNA or RNA encoding a human nNR5 protein. Compounds which modulate these activities may be DNA, RNA, peptides, proteins, or non-proteinaceous organic molecules. Compounds may modulate by increasing or attenuating the expression of DNA or RNA encoding human nNR5, or the function of human nNR5. Compounds that modulate the expression of DNA or RNA encoding human nNR5 or the biological function thereof may be detected by a variety of assays. The assay may be a simple "yes/no" assay to determine whether there is a change in expression or function. The assay may be made quantitative by comparing the expression or function of a test sample with the levels of expression or function in a standard sample. Kits containing human nNR5, antibodies to human nNR5, or modified human nNR5 may be prepared by known methods for such uses.

The DNA molecules, RNA molecules, recombinant protein and antibodies of the present invention may be used to screen and measure levels of human nNR5. The recombinant proteins, DNA molecules, RNA molecules and antibodies lend themselves to the formulation of kits suitable for the detection and typing of human nNR5. Such a kit would comprise a compartmentalized carrier suitable to hold in close confinement at least one container. The carrier would further comprise reagents such as recombinant nNR5 or anti-nNR5 antibodies suitable for detecting human nNR5. The carrier may also contain a means for detection such as labeled antigen or enzyme substrates or the like.

Pharmaceutically useful compositions comprising modulators of human nNR5 may be formulated according to known methods such as by the admixture of a pharmaceutically acceptable carrier. Examples of such carriers and methods of formulation may be found in Remington's Pharmaceutical Sciences. To form a pharmaceutically acceptable composition suitable for effective administration, such compositions will contain an effective amount of

the protein, DNA, RNA, modified human nNR5, or either nNR5 agonsits or antagonists.

Therapeutic or diagnostic compositions comprising modulators of nNR5 are administered to an individual in amounts sufficient to treat or diagnose disorders. The effective amount may vary according to a variety of factors such as the individual's condition, weight, sex and age. Other factors include the mode of administration.

The pharmaceutical compositions may be provided to the individual by a variety of routes such as subcutaneous, topical, oral and intramuscular.

The term "chemical derivative" describes a molecule that contains additional chemical moieties which are not normally a part of the base molecule. Such moieties may improve the solubility, half-life, absorption, etc. of the base molecule. Alternatively the moieties may attenuate undesirable side effects of the base molecule or decrease the toxicity of the base molecule. Examples of such moieties are described in a variety of texts, such as Remington's Pharmaceutical Sciences.

Compounds identified according to the methods disclosed herein may be used alone at appropriate dosages. Alternatively, co-administration or sequential administration of other agents may be desirable.

The present invention also has the objective of providing suitable topical, oral, systemic and parenteral pharmaceutical formulations for use in the novel methods of treatment of the present invention. The compositions containing compounds identified according to this invention as the active ingredient can be administered in a wide variety of therapeutic dosage forms in conventional vehicles for administration. For example, the compounds can be administered in such oral dosage forms as tablets, capsules (each including timed release and sustained release formulations), pills, powders, granules, elixirs, tinctures, solutions, suspensions, syrups and emulsions, or by injection. Likewise, they may also be administered in intravenous (both bolus and infusion), intraperitoneal, subcutaneous, topical with or without occlusion, or intramuscular form, all using forms well known to those of ordinary skill in the pharmaceutical arts.

Advantageously, compounds of the present invention may be administered in a single daily dose, or the total daily dosage may be administered in divided doses of two, three or four times daily. Furthermore, compounds for the present invention can be administered in intranasal form via topical use of suitable intranasal vehicles, or via transdermal routes, using those forms of transdermal skin patches well known to those of ordinary skill in that art. To be administered in the form of a transdermal delivery system, the dosage administration will, of course, be continuous rather than intermittent throughout the dosage regimen.

For combination treatment with more than one active agent, where the active agents are in separate dosage formulations, the active agents can be administered concurrently, or they each can be administered at separately staggered times.

The dosage regimen utilizing the compounds of the present invention is selected in accordance with a variety of factors including type, species, age, weight, sex and medical condition of the patient; the severity of the condition to be treated; the route of administration; the renal, hepatic and cardiovascular function of the patient; and the particular compound thereof employed. A physician or veterinarian of ordinary skill can readily determine and prescribe the effective amount of the drug required to prevent, counter or arrest the progress of the condition. Optimal precision in achieving concentrations of drug within the range that yields efficacy without toxicity requires a regimen based on the kinetics of the drug's availability to target sites. This involves a consideration of the distribution, equilibrium, and elimination of a drug.

The following examples are provided to illustrate the present invention without, however, limiting the same hereto.

EXAMPLE 1:

Isolation and Characterization of a DNA Molecule Encoding nNR5

An EST from a human retina cDNA library was identified during a data base search. This EST is identified by GenBank Accession No. W27871 and dbEST Id No. 534939 and is disclosed as follows:

```
1    GGAATCACCA GGGGAGACAG GNGCACAGNG AGACAGAGGT TCATGGACTG
51   AGGCAAAGGC TGGGCCAGGC TCAGCAACCC AGGCCTCCCG CAGGCAGGCA
101  GAGGCTGCCC TGTAACCCAT GGAGACCAGA CCAACAGCTC TGATGAGCTC
151  CACAGTGGCT GCAGCTGCGC CTGCAGCTGG GGCTGCCTCC AGGAAGGAGT
201  CTCCAGGCAG ATGGGGCCTG GGGGAGGATC CCACAGGCGT GAGCCCCTCG
251  CTCCAGTGCC GCGTGTGCGG AGACAGCAGC AGCGGGAAGC ACTATGGCAT
301  CTATGCCCTG CAACGGTTGC AGCGGTTTCT TCCAAGAGGA GCNGTACGGN
351  GGAGGCTCAA TCCTTACAAG GGTGCCCAGG GTGGGGGCAG GGATTGTGCC
401  CCCCNGTGGA CAAGGNCCCA ACCCGNAACC CAGTGCCCAG GCCTGCCGGN
451  TTGAGAAGTG CTTNAAAANN NGGNNGGGGN TTGAACCCAG GACGCCCGTN
501  NAAAGGAACG ANNGCCNAGC CCGNGAGGAN AAGCCCAGGT NCCACCCCTG
551  GANAAGAATN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
601  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
651  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
701  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
751  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
801  NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
851  NNNNNNNNNN (SEQ ID NO:3).
```

DNA fragments encoding DBD regions of androgen receptor (AR), estrogen receptor b (ERb), glucocorticoid receptor (GR) and vitamin D receptor (VDR) were generated by PCR and subcloned into pCR cloning vectors as described by the manufacturer. The following oligonucleotide primers were utilized to generate fragments for plasmid subcloning:

1. GR-R 5'-TTTCGAGCTTCCAGGTTCAT-3' (SEQ ID NO: 6),
2. GR-F 5'-CTCCCAAACTCTGCCTGGTG-3' (SEQ ID NO: 7),
3. ERB-R 5'-CGGGAGCCACACTTCACCAT-3' (SEQ ID NO: 8),

4. ERB-F 5'-GCTCACTTCTGCGCTGTCTG-3' (SEQ ID NO: 9),
5. AR-R 5'-TTCCGGGCTCCCAGAGTCAT-3' (SEQ ID NO: 10),
6. AR-F 5'-CAGAAGACCTGCCTGATCTG-3' (SEQ ID NO:11),
7. VDR-R 5'-GAAATGAACTCCTTCATCAT-3' (SEQ ID NO: 12),
8. VDR-F 5'-CCGGATCTGTGGGGTGTGTG-3' (SEQ ID NO: 13).

PCR templates for AR, ERb and GR are cDNAs made from human fetal brain mRNA. PCR template for VDR was a cDNA made from human small intestine mRNA. The DNA fragments were purified using a Qiagen gel extraction kit. Phosphorylation, self-ligation and transformation of the purified DNA was carried out as recommended by the manufacturer. A human retina cDNA library was screened at low stringency using the above-identified AR, Erb, GR and VDR's DBD regions as probes. Two positive clones were selected and subjected to sequence analysis, which revealed the presence of an intron as shown herein and as set forth as SEQ ID NO: 18. Direct sequencing of plasmid DNA from clone A8 and A9 revealed a full cDNA molecule 3,012 bps in length (SEQ ID NO: 18), which encodes a peptide most related to hCOUP-TF (Wang et al., 1989, *Nature* 340: 163-166). These cDNA clones showed homology to the human EST (GenBank Accession No. W27871 and dbEST Id No. 534939; SEQ ID NO: 3).

To isolate an intronless cDNA clone for nNR5, the human retina cDNA library was screened by PCR analysis with primer pair nNR5F2 (5'-ATGAGCTCCACAGTGGCTGC-3'; SEQ ID NO: 4) and nNR5R (5'-CTGTCTCCGCACACGCGGCA-3'; SEQ ID NO: 5) from the human EST (GenBank Accession No. W27871 and dbEST Id No. 534939; SEQ ID NO: 3). Further screening of the retina cDNA library by PCR using nNR5F2/nNR5R on retina cDNA resulted in a total of 20 positive clones from approximately 250,000 primary clones. This data indicated that the gene of interest (eventually identified as a cDNA encoding human nNR5) is abundantly expressed in retina tissue. In order to define the exact intron-exon boundary and to isolate an intronless cDNA, primer pair R5F3 (5'-CTGATGAGAATATTGATGT-3'; SEQ ID NO: 14) and R5R4 (5'-CGTGAGCCGGCCCTGGGCA-3'; SEQ ID NO: 15), which flank the putative intron region, was used in PCR on the twenty positive clones. Two clones, E1 and F6, yielded a band of smaller size than that

of the A8 which had an intron. DNA fragments from this PCR were purified and submitted for sequencing. Automated sequencing was performed on and sequence assembly and analysis were performed with SEQUENCHER™ 3.0 (Gene Codes Corporation, Ann Arbor, MI). Ambiguities and/or discrepancies between automated base calling in sequencing reads were visually examined and edited to the correct base call. Based on the sequencing result and protein sequence alignment an intron region in the original A8/A9 clone was identified from nucleotide 971 to 1847. Therefore, the full length cDNA without an intron is approximately 2.1kb and this DNA molecule which encodes human nNR5 is shown in Figure 1A-B and is set forth as SEQ ID NO: 1.

In order to identify the genome map position of nNR5, primers in the 3' non-coding region were designed. Forward primer R5F9 (5'-GGCATGGACCTCACTGAAGA-3'; SEQ ID NO: 16) and reverse primer R5R10 (5'-ACTGGCAGGAACCTGTTATA-3'; SEQ ID NO: 17) were used in PCR scanning on the 83 clones of the Stanford radiation hybrid panel (Cox et al., 1990, *Science*, 250:245-250). The PCR results were scored and submitted to the Stanford Genome Center for linkage analysis. The result indicate that nNR5 is located on chromosome 15.

WHAT IS CLAIMED:

1. A purified DNA molecule encoding a human nNR5 protein wherein said protein comprises the amino acid sequence as follows:

METRPTALMS STVAAAAPAA GAASRKESPG RWGLGEDPTG VSPSLQCRVC
GDSSSGKHYG IYACNGCSGF FKRSVRRRLI YRCQVGAGMC PVDKAHRNQC
QACRLKKCLQ AGMNQDAVQN ERQPRSTAQV HLDSMESNTE SRPESLVAPP
APAGRSPRGP TPMSAARALG HHFMASLITA ETCAKLEPED ADENIDVTSN
DPEFPSSPYS SSSPCGLDSI HETSARLLFM AVKWAKNLPV FSSLPFRDQV
ILLEEAWSEL FLLGAIQWSL PLDSCPLLAP PEASAAGGAQ GRLTLASMET
RVLQETISRF RALAVDPTEF ACMKALVLFK PETRGLKDPE HVEALQDQSQ
VMLSQHSKAH HPSQPVR, as set forth in three-letter abbreviation in SEQ ID NO:2.

2. An expression vector for expressing a human nNR5 protein in a recombinant host cell wherein said expression vector comprises a DNA molecule of claim 1.

3. A host cell which expresses a recombinant human nNR5 protein wherein said host cell contains the expression vector of claim 2.

4. A process for expressing a human nNR5 protein in a recombinant host cell, comprising:

(a) transfecting the expression vector of claim 2 into a suitable host cell; and,

(b) culturing the host cells of step (a) under conditions which allow expression of said the human nNR5 protein from said expression vector.

5. A purified DNA molecule encoding a human nNR5 protein wherein said protein consists of the amino acid sequence as follows:

METRPTALMS STVAAAAPAA GAASRKESPG RWGLGEDPTG VSPSLQCRVC
GDSSSGKHYG IYACNGCSGF FKRSVRRRLI YRCQVGAGMC PVDKAHRNQC
QACRLKKCLQ AGMNQDAVQN ERQPRSTAQV HLDSMESNTE SRPESLVAPP
APAGRSPRGP TPMSAARALG HHFMASLITA ETCAKLEPED ADENIDVTSN
DPEFPSSPYS SSSPCGLDSI HETSARLLFM AVKWAKNLPV FSSLPFRDQV
ILLEEAWSEL FLLGAIQWSL PLDSCPLLAP PEASAAGGAQ GRLTLASMET
RVLQETISRF RALAVDPTEF ACMKALVLFK PETRGLKDPE HVEALQDQSQ
VMLSQHSKAH HPSQPVR,as set forth in three-letter abbreviation in SEQ ID NO:2.

6. An expression vector for expressing a human nNR5 protein in a recombinant host cell wherein said expression vector comprises a DNA molecule of claim 5.

7. A host cell which expresses a recombinant human nNR5 protein wherein said host cell contains the expression vector of claim 6.

8. A process for expressing a human nNR5 protein in a recombinant host cell, comprising:

(a) transfecting the expression vector of claim 6 into a suitable host cell; and,

(b) culturing the host cells of step (a) under conditions which allow expression of said the human nNR5 protein from said expression vector.

9. A purified DNA molecule encoding a human nNR5 protein wherein said DNA molecule comprises the nucleotide sequence as set forth in SEQ ID NO: 1, as follows:

ATTCGGGACC TTGGGGCAGC TCCTGAGTTC AGACAGAGTT CAGGAAGGGA
GACAGGGGCA CAGAGAGACA GAGGTTCATG GACTGAGGCA AAGGCTGGGC
CAGGCTCAGC AACCCAGGCC TCCCGCAGGC AGGCAGAGGC TGCCCTGTAA
CCCATGGAGA CCAGACCAAC AGCTCTGATG AGCTCCACAG TGGCTGCAGC

TGCGCCTGCA GCTGGGGCTG CCTCCAGGAA GGAGTCTCCA GGCAGATGGG
GCCTGGGGGA GGATCCCACA GGCGTGAGCC CCTCGCTCCA GTGCCGCGTG
TGCGGAGACA GCAGCAGCGG GAAGCACTAT GGCATCTATG CCTGCAACGG
CTGCAGCGGC TTCTTCAAGA GGAGCGTACG GCGGAGGCTC ATCTACAGGT
GCCAGGTGGG GGCAGGGATG TGCCCCGTGG ACAAGGCCCA CCGCAACCAG
TGCCAGGCCT GCCGGCTGAA GAAGTGCCTG CAGGCGGGGA TGAACCAGGA
CGCCGTGCAG AACGAGCGCC AGCCGCGAAG CACAGCCCAG GTCCACCTGG
ACAGCATGGA GTCCAACACT GAGTCCCGGC CGGAGTCCCT GGTGGCTCCC
CCGGCCCCGG CAGGGCGCAG CCCACGGGGC CCCACACCCA TGTCTGCAGC
CAGAGCCCTG GGCCACCACT TCATGGCCAG CCTTATAACA GCTGAAACCT
GTGCTAAGCT GGAGCCAGAG GATGCTGATG AGAATATTGA TGTCACCAGC
AATGACCCTG AGTTCCCCTC CTCTCCATAC TCCTCTTCCT CCCCCTGCGG
CCTGGACAGC ATCCATGAGA CCTCGGCTCG CCTACTCTTC ATGGCCGTCA
AGTGGGCCAA GAACCTGCCT GTGTTCTCCA GCCTGCCCTT CCGGGATCAG
GTGATCCTGC TGGAAGAGGC GTGGAGTGAA CTCTTTCTCC TCGGGGCCAT
CCAGTGGTCT CTGCCTCTGG ACAGCTGTCC TCTGCTGGCA CCGCCCGAGG
CTTCTGCTGC CGGTGGTGCC CAGGGCCGGC TCACGCTGGC CAGCATGGAG
ACGCGTGTCC TGCAGGAAAC TATCTCTCGG TTCCGGGCAT TGGCGGTGGA
CCCCACGGAG TTTGCCTGCA TGAAGGCCTT GGTCCTCTTC AAGCCAGAGA
CGCGGGGCCT GAAGGATCCT GAGCACGTAG AGGCCTTGCA GGACCAGTCC
CAAGTGATGC TGAGCCAGCA CAGCAAGGCC CACCACCCCA GCCAGCCCGT
GAGGTGACCT GAGCATGCGC CCACCCACTC ATCTGTCCCT GACCTCTAAC
CTTTCTCTGC CTCTCCCACA CTCTCCCAGA GCTCACTGAT TAGACAGCAC
AAGGGTCTCA GTTCAACAGC ATACAGCCAA CATCTATGGT GTCCCAGGCA
CAGTGCCAGG CCCCGGGAGT GGGGACCAAG ATGTACATAA GACAAAGCTA
CTGCCTTCTA GAGACAACCG GCAGTGACCT CACTGAAGAC AAAAACTGCC
CTAGCCAGGT ACTGAGGGTT GCATGAATCT GCAGGAGACA GAGATCCCCT
TGCATGGGAA ACATAAAGCA GAATTGGGAG GGACTTTGTG GAGACAGGGC
TGGACTTGAA AGGAAGAAGA AGTCTAAAAG AAAACATCAT TTGCAAAGGG
AGAGAGGGGC AAGCATGATA TGTTGTTAGA ACAGGAGCCC ACTTTGAAGG
TATAACAGGT TCCTGCCAGT GAGAAATGGG GAGAATAAGC CAGAAAAGTA
CCCTAGGACC AGCCCGTTCA GGACTTTGAA TGCCAGCCAA AGGCCACGTC
TGACTTGGGA GGCAGAGGGC AGCTACTGCA GGTTTCCGAG CAGAGGGTCA
TACACAGGGC TGGACCTCAC GCAGACTGGC ATGGCCATGG GTCCAGAGGA
TACTACTGGG AAGGGGATGG CAGCTACTGC CACCTTCCAG ATGGTTCCAT

GGAGTTCTGA TCTTTGGGCA TGGCCAGGGG AAGCAGAAGG GAGACTCTAG GAGTTGAAAT GGGTCAGACC CGGTGTTTGG GTGAAGGTAA GGAATGAGGG AAGAGGAGCT CTTTG (SEQ ID NO: 1).

10. A DNA molecule of claim 9 which consists of nucleotide 154 to about nucleotide 1257 of SEQ ID NO: 1.

11. An expression vector for expressing a human nNR5 protein wherein said expression vector comprises a DNA molecule of claim 9.

12. An expression vector for expressing a human nNR5 protein wherein said expression vector comprises a DNA molecule of claim 11.

13. A host cell which expresses a recombinant human nNR5 protein wherein said host cell contains the expression vector of claim 11.

14. A host cell which expresses a recombinant human nNR5 protein wherein said host cell contains the expression vector of claim 12.

15. A process for expressing a human nNR5 protein in a recombinant host cell, comprising:

(a) transfecting the expression vector of claim 11 into a suitable host cell; and,

(b) culturing the host cells of step (a) under conditions which allow expression of said the human nNR5 protein from said expression vector.

16. A purified DNA molecule encoding a human nNR5 protein wherein said DNA molecule consists of the nucleotide sequence as set forth in SEQ ID NO: 1, as follows:

ATTCGGGACC TTGGGGCAGC TCCTGAGTTC AGACAGAGTT CAGGAAGGGA
GACAGGGGCA CAGAGAGACA GAGGTTCATG GACTGAGGCA AAGGCTGGGC
CAGGCTCAGC AACCCAGGCC TCCCGCAGGC AGGCAGAGGC TGCCCTGTAA
CCCATGGAGA CCAGACCAAC AGCTCTGATG AGCTCCACAG TGGCTGCAGC
TGCGCCTGCA GCTGGGGCTG CCTCCAGGAA GGAGTCTCCA GGCAGATGGG
GCCTGGGGGA GGATCCCACA GGCGTGAGCC CCTCGCTCCA GTGCCGCGTG
TGCGGAGACA GCAGCAGCGG GAAGCACTAT GGCATCTATG CCTGCAACGG
CTGCAGCGGC TTCTTCAAGA GGAGCGTACG GCGGAGGCTC ATCTACAGGT
GCCAGGTGGG GGCAGGGATG TGCCCCGTGG ACAAGGCCCA CCGCAACCAG
TGCCAGGCCT GCCGGCTGAA GAAGTGCCTG CAGGCGGGGA TGAACCAGGA
CGCCGTGCAG AACGAGCGCC AGCCGCGAAG CACAGCCCAG GTCCACCTGG
ACAGCATGGA GTCCAACACT GAGTCCCGGC CGGAGTCCCT GGTGGCTCCC
CCGGCCCCGG CAGGGCGCAG CCCACGGGGC CCCACACCCA TGTCTGCAGC
CAGAGCCCTG GGCCACCACT TCATGGCCAG CCTTATAACA GCTGAAACCT
GTGCTAAGCT GGAGCCAGAG GATGCTGATG AGAATATTGA TGTCACCAGC
AATGACCCTG AGTTCCCCTC CTCTCCATAC TCCTCTTCCT CCCCCTGCGG
CCTGGACAGC ATCCATGAGA CCTCGGCTCG CCTACTCTTC ATGGCCGTCA
AGTGGGCCAA GAACCTGCCT GTGTTCTCCA GCCTGCCCTT CCGGGATCAG
GTGATCCTGC TGGAAGAGGC GTGGAGTGAA CTCTTTCTCC TCGGGGCCAT
CCAGTGGTCT CTGCCTCTGG ACAGCTGTCC TCTGCTGGCA CCGCCCGAGG
CTTCTGCTGC CGGTGGTGCC CAGGGCCGGC TCACGCTGGC CAGCATGGAG
ACGCGTGTCC TGCAGGAAAC TATCTCTCGG TTCCGGGCAT TGGCGGTGGA
CCCCACGGAG TTTGCCTGCA TGAAGGCCTT GGTCCTCTTC AAGCCAGAGA
CGCGGGGCCT GAAGGATCCT GAGCACGTAG AGGCCTTGCA GGACCAGTCC
CAAGTGATGC TGAGCCAGCA CAGCAAGGCC CACCACCCCA GCCAGCCCGT
GAGGTGACCT GAGCATGCGC CCACCCACTC ATCTGTCCCT GACCTCTAAC
CTTTCTCTGC CTCTCCCACA CTCTCCCAGA GCTCACTGAT TAGACAGCAC
AAGGGTCTCA GTTCAACAGC ATACAGCCAA CATCTATGGT GTCCCAGGCA
CAGTGCCAGG CCCCGGGAGT GGGGACCAAG ATGTACATAA GACAAAGCTA
CTGCCTTCTA GAGACAACCG GCAGTGACCT CACTGAAGAC AAAAACTGCC
CTAGCCAGGT ACTGAGGGTT GCATGAATCT GCAGGAGACA GAGATCCCCT
TGCATGGGAA ACATAAAGCA GAATTGGGAG GGACTTTGTG GAGACAGGGC

TGGACTTGAA AGGAAGAAGA AGTCTAAAAG AAAACATCAT TTGCAAAGGG
AGAGAGGGGC AAGCATGATA TGTTGTTAGA ACAGGAGCCC ACTTTGAAGG
TATAACAGGT TCCTGCCAGT GAGAAATGGG GAGAATAAGC CAGAAAAGTA
CCCTAGGACC AGCCCGTTCA GGACTTTGAA TGCCAGCCAA AGGCCACGTC
TGACTTGGGA GGCAGAGGGC AGCTACTGCA GGTTTCCGAG CAGAGGGTCA
TACACAGGGC TGGACCTCAC GCAGACTGGC ATGGCCATGG GTCCAGAGGA
TACTACTGGG AAGGGGATGG CAGCTACTGC CACCTTCCAG ATGGTTCCAT
GGAGTTCTGA TCTTTGGGCA TGGCCAGGGG AAGCAGAAGG GAGACTCTAG
GAGTTGAAAT GGGTCAGACC CGGTGTTTGG GTGAAGGTAA GGAATGAGGG
AAGAGGAGCT CTTTG (SEQ ID NO: 1).

17. A DNA molecule of claim 16 which consists of nucleotide 154 to about nucleotide 1257 of SEQ ID NO: 1.

18. An expression vector for expressing a human nNR5 protein wherein said expression vector comprises a DNA molecule of claim 16.

19. An expression vector for expressing a human nNR5 protein wherein said expression vector comprises a DNA molecule of claim 17.

20. A host cell which expresses a recombinant human nNR5 protein wherein said host cell contains the expression vector of claim 18.

21. A host cell which expresses a recombinant human nNR5 protein wherein said host cell contains the expression vector of claim 19.

22. A process for expressing a human nNR5 protein in a recombinant host cell, comprising:

(a) transfecting the expression vector of claim 18 into a suitable host cell; and,

(b) culturing the host cells of step (a) under conditions which allow expression of said the human nNR5 protein from said expression vector.

23. A purified DNA molecule encoding a human nNR5 protein wherein said DNA molecule comprises the nucleotide sequence as set forth in SEQ ID NO: 19, as follows:

TATAGGGCGA ATTGGGTACC GGGCCCCCCC TCGAGGTCGA CGGTATCGAT
AAGCTTGATA TCGAATTCGA ATTCGGGACC TTGGGGCAGC TCCTGAGTTC
AGACAGAGTT CAGGAAGGGA GACAGGGGCA CAGAGAGACA GAGGTTCATG
GACTGAGGCA AAGGCTGGGC CAGGCTCAGC AACCCAGGCC TCCCGCAGGC
AGGCAGAGGC TGCCCTGTAA CCCATGGAGA CCAGACCAAC AGCTCTGATG
AGCTCCACAG TGGCTGCAGC TGCGCCTGCA GCTGGGGCTG CCTCCAGGAA
GGAGTCTCCA GGCAGATGGG GCCTGGGGGA GGATCCCACA GGCGTGAGCC
CCTCGCTCCA GTGCCGCGTG TGCGGAGACA GCAGCAGCGG GAAGCACTAT
GGCATCTATG CCTGCAACGG CTGCAGCGGC TTCTTCAAGA GGAGCGTACG
GCGGAGGCTC ATCTACAGGT GCCAGGTGGG GGCAGGGATG TGCCCCGTGG
ACAAGGCCCA CCGCAACCAG TGCCAGGCCT GCCGGCTGAA GAAGTGCCTG
CAGGCGGGGA TGAACCAGGA CGCCGTGCAG AACGAGCGCC AGCCGCGAAG
CACAGCCCAG GTCCACCTGG ACAGCATGGA GTCCAACACT GAGTCCCGGC
CGGAGTCCCT GGTGGCTCCC CCGGCCCCGG CAGGGCGCAG CCCACGGGGC
CCCACACCCA TGTCTGCAGC CAGAGCCCTG GGCCACCACT TCATGGCCAG
CCTTATAACA GCTGAAACCT GTGCTAAGCT GGAGCCAGAG GATGCTGATG
AGAATATTGA TGTCACCAGC AATGACCCTG AGTTCCCCTC CTCTCCATAC
TCCTCTTCCT CCCCCTGCGG CCTGGACAGC ATCCATGAGA CCTCGGCTCG
CCTACTCTTC ATGGCCGTCA AGTGGGCCAA GAACCTGCCT GTGTTCTCCA
GCCTGCCCTT CCGGGATCAG GTGATCCTGC TGGAAGAGGC GTGGAGTGAA
CTCTTTCTCC TCGGGGCCAT CCAGTGGTCT CTGCCTCTGG ACAGCTGTCC
TCTGCTGGCA CCGCCCGAGG CCTCTGCTGC CGGTGGTGCC CAGGGCCGGC
TCACGCTGGC CAGCATGGAG ACGCGTGTCC TGCAGGAAAC TATCTCTCGG
TTCCGGGCAT TGGCGGTGGA CCCCACGGAG TTTGCCTGCA TGAAGGCCTT
GGTCCTCTTC AAGCCAGAGA CGCGGGGCCT GAAGGATCCT GAGCACGTAG
AGGCCTTGCA GGACCAGTCC CAAGTGATGC TGAGCCAGCA CAGCAAGGCC
CACCACCCCA GCCAGCCCGT GAGGTGACCT GAGCATGCGC CCACCCACTC

ATCTGTCCCT GACCTCTAAC CTTTCTCTGC CTCTCCCACA CTCTCCCAGA
GCTCACTGAT TAGACAGCAC AAGGGTCTCA GTTCAACAGC ATACAGCCAA
CATCTATGGT GTCCCAGGCA CAGTGCCAGG CCCCGGGAGT GGGGACCAAG
ATGTACATAA GACAAAGCTA CTGCCTTCTA GAGACAACCG GCAGTGACCT
CACTGAAGAC AAAAACTGCC CTAGCCAGGT ACTGAGGGTT GCATGAATCT
GCAGGAGACA GAGATCCCCT TGCATGGGAA ACATAAAGCA GAATTGGGAG
GGACTTTGTG GAGACAGGGC TGGACTTGAA AGGAAGAAGA AGTCTAAAAG
AAAACATCAT TTGCAAAGGG AGAGAGGGGC AAGCATGATA TGTTGTTAGA
ACAGGAGCCC ACTTTGAAGG TATAACAGGT TCCTGCCAGT GAGAAATGGG
GAGAATAAGC CAGAAAAGTA CCCTAGGACC AGCCCGTTCA GGACTTTGAA
TGCCAGCCAA AGGCCACGTC TGACTTGGGA GGCAGAGGGC AGCTACTGCA
GGTTTCCGAG CAGAGGGTCA TACACAGGGC TGGACCTCAC GCAGACTGGC
ATGGCCATGG GTCCAGAGGA TACTACTGGG AAGGGGATGG CAGCTACTGC
CACCTTCCAG ATGGTTCCAT GGAGTTCTGA TCTTTGGGCA TGGCCAGGGG
AAGCAGAAGG GAGACTCTAG GAGTTGAAAT GGGTCAGACC CGGTGTTTGG
GTGAAGGTAA GGAATGAGGG AAGAGGAGCT CTTTG (SEQ ID NO: 19).

24. An expression vector for expressing a human nNR5 protein wherein said expression vector comprises a DNA molecule of claim 23.

25. A host cell which expresses a recombinant human nNR5 protein wherein said host cell contains the expression vector of claim 24.

26. A process for expressing a human nNR5 protein in a recombinant host cell, comprising:

(a) transfecting the expression vector of claim 24 into a suitable host cell; and,

(b) culturing the host cells of step (a) under conditions which allow expression of said the human nNR5 protein from said expression vector.

27. A DNA molecule of claim 23 which consists of nucleotide 224 to about nucleotide 1327 of SEQ ID NO: 19.

28. A purified human nNR5 protein which comprises the amino acid sequence as set forth in SEQ ID NO: 2.

29. The purified human nNR5 protein of claim 28 which consists of the amino acid sequence as set forth in SEQ ID NO: 2.

1 ATTCGGGACC TTGGGGCAGC TCCTGAGTTC AGACAGAGTT CAGGAAGGGA

51 GACAGGGGCA CAGAGAGACA GAGGTTCATG GACTGAGGCA AAGGCTGGGC

101 CAGGCTCAGC AACCCAGGCC TCCCGCAGGC AGGCAGAGGC TGCCCTGTAA

151 CCCATGGAGA CCAGACCAAC AGCTCTGATG AGCTCCACAG TGGCTGCAGC

201 TGCGCCTGCA GCTGGGGCTG CCTCCAGGAA GGAGTCTCCA GGCAGATGGG

251 GCCTGGGGGA GGATCCCACA GGCGTGAGCC CCTCGCTCCA GTGCCGCGTG

301 TGCGGAGACA GCAGCAGCGG GAAGCACTAT GGCATCTATG CCTGCAACGG

351 CTGCAGCGGC TTCTTCAAGA GGAGCGTACG GCGGAGGCTC ATCTACAGGT

401 GCCAGGTGGG GGCAGGGATG TGCCCCGTGG ACAAGGCCCA CCGCAACCAG

451 TGCCAGGCCT GCCGGCTGAA GAAGTGCCTG CAGGCGGGGA TGAACCAGGA

501 CGCCGTGCAG AACGAGCGCC AGCCGCGAAG CACAGCCCAG GTCCACCTGG

551 ACAGCATGGA GTCCAACACT GAGTCCCGGC CGGAGTCCCT GGTGGCTCCC

601 CCGGCCCCGG CAGGGCGCAG CCCACGGGGC CCCACACCCA TGTCTGCAGC

651 CAGAGCCCTG GGCCACCACT TCATGGCCAG CCTTATAACA GCTGAAACCT

701 GTGCTAAGCT GGAGCCAGAG GATGCTGATG AGAATATTGA TGTCACCAGC

751 AATGACCCTG AGTTCCCCTC CTCTCCATAC TCCTCTTCCT CCCCCTGCGG

801 CCTGGACAGC ATCCATGAGA CCTCGGCTCG CCTACTCTTC ATGGCCGTCA

851 AGTGGGCCAA GAACCTGCCT GTGTTCTCCA GCCTGCCCTT CCGGGATCAG

901 GTGATCCTGC TGGAAGAGGC GTGGAGTGAA CTCTTTCTCC TCGGGGCCAT

951 CCAGTGGTCT CTGCCTCTGG ACAGCTGTCC TCTGCTGGCA CCGCCCGAGG

FIG.1A

1001 CTTCTGCTGC CGGTGGTGCC CAGGGCCGGC TCACGCTGGC CAGCATGGAG

1051 ACGCGTGTCC TGCAGGAAAC TATCTCTCGG TTCCGGGCAT TGGCGGTGGA

1101 CCCCACGGAG TTTGCCTGCA TGAAGGCCTT GGTCCTCTTC AAGCCAGAGA

1151 CGCGGGGCCT GAAGGATCCT GAGCACGTAG AGGCCTTGCA GGACCAGTCC

1201 CAAGTGATGC TGAGCCAGCA CAGCAAGGCC CACCACCCCA GCCAGCCCGT

1251 GAGGTGACCT GAGCATGCGC CCACCCACTC ATCTGTCCCT GACCTCTAAC

1301 CTTTCTCTGC CTCTCCCACA CTCTCCCAGA GCTCACTGAT TAGACAGCAC

1351 AAGGGTCTCA GTTCAACAGC ATACAGCCAA CATCTATGGT GTCCCAGGCA

1401 CAGTGCCAGG CCCCGGGAGT GGGGACCAAG ATGTACATAA GACAAAGCTA

1451 CTGCCTTCTA GAGACAACCG GCAGTGACCT CACTGAAGAC AAAAACTGCC

1501 CTAGCCAGGT ACTGAGGGTT GCATGAATCT GCAGGAGACA GAGATCCCCT

1551 TGCATGGGAA ACATAAAGCA GAATTGGGAG GGACTTTGTG GAGACAGGGC

1601 TGGACTTGAA AGGAAGAAGA AGTCTAAAAG AAAACATCAT TTGCAAAGGG

1651 AGAGAGGGGC AAGCATGATA TGTTGTTAGA ACAGGAGCCC ACTTTGAAGG

1701 TATAACAGGT TCCTGCCAGT GAGAAATGGG GAGAATAAGC CAGAAAAGTA

1751 CCCTAGGACC AGCCCGTTCA GGACTTTGAA TGCCAGCCAA AGGCCACGTC

1801 TGACTTGGGA GGCAGAGGGC AGCTACTGCA GGTTTCCGAG CAGAGGGTCA

1851 TACACAGGGC TGGACCTCAC GCAGACTGGC ATGGCCATGG GTCCAGAGGA

1901 TACTACTGGG AAGGGGATGG CAGCTACTGC CACCTTCCAG ATGGTTCCAT

1951 GGAGTTCTGA TCTTTGGGCA TGGCCAGGGG AAGCAGAAGG GAGACTCTAG

2001 GAGTTGAAAT GGGTCAGACC CGGTGTTTGG GTGAAGGTAA GGAATGAGGG

2051 AAGAGGAGCT CTTTG (SEQ ID NO:1)

FIG. 1B

```
   1 ATTCGGGACCTTGGGGCAGCTCCTGAGTTCAGACAGAGTTCAGGAAGGGAGACAGGGGCA  60

  61 CAGAGAGACAGAGGTTCATGGACTGAGGCAAAGGCTGGGCCAGGCTCAGCAACCCAGGCC 120

 121 TCCCGCAGGCAGGCAGAGGCTGCCCTGTAACCCATGGAGACCAGACCAACAGCTCTGATG
                                     M  E  T  R  P  T  A  L  M

 181 AGCTCCACAGTGGCTGCAGCTGCGCCTGCAGCTGGGGCTGCCTCCAGGAAGGAGTCTCCA 240
     S  S  T  V  A  A  A  A  P  A  A  G  A  A  S  R  K  E  S  P

 241 GGCAGATGGGGCCTGGGGGAGGATCCCACAGGCGTGAGCCCCTCGCTCCAGTGCCGCGTG 300
     G  R  W  G  L  G  E  D  P  T  G  V  S  P  S  L  Q  C  R  V

 301 TGCGGAGACAGCAGCAGCGGGAAGCACTATGGCATCTATGCCTGCAACGGCTGCAGCGGC 360
     C  G  D  S  S  S  G  K  H  Y  G  I  Y  A  C  N  G  C  S  G

 361 TTCTTCAAGAGGAGCGTACGGCGGAGGCTCATCTACAGGTGCCAGGTGGGGGCAGGGATG 420
     F  F  K  R  S  V  R  R  R  L  I  Y  R  C  Q  V  G  A  G  M

 421 TGCCCCGTGGACAAGGCCCACCGCAACCAGTGCCAGGCCTGCCGGCTGAAGAAGTGCCTG 480
     C  P  V  D  K  A  H  R  N  Q  C  Q  A  C  R  L  K  K  C  L

 481 CAGGCGGGGATGAACCAGGACGCCGTGCAGAACGAGCGCCAGCCGCGAAGCACAGCCCAG 540
     Q  A  G  M  N  Q  D  A  V  Q  N  E  R  Q  P  R  S  T  A  Q

 541 GTCCACCTGGACAGCATGGAGTCCAACACTGAGTCCCGGCCGGAGTCCCTGGTGGCTCCC 600
     V  H  L  D  S  M  E  S  N  T  E  S  R  P  E  S  L  V  A  P

 601 CCGGCCCCGGCAGGGCGCAGCCCACGGGGCCCCACACCCATGTCTGCAGCCAGAGCCCTG 660
     P  A  P  A  G  R  S  P  R  G  P  T  P  M  S  A  A  R  A  L

 661 GGCCACCACTTCATGGCCAGCCTTATAACAGCTGAAACCTGTGCTAAGCTGGAGCCAGAG 720
     G  H  H  F  M  A  S  L  I  T  A  E  T  C  A  K  L  E  P  E

 721 GATGCTGATGAGAATATTGATGTCACCAGCAATGACCCTGAGTTCCCCTCCTCTCCATAC 780
     D  A  D  E  N  I  D  V  T  S  N  D  P  E  F  P  S  S  P  Y

 781 5CCTCTTCCTCCCCCTGCGGCCTGGACAGCATCCATGAGACCTCGGCTCGCCTACTCTTC 840
     S  S  S  S  P  C  G  L  D  S  I  H  E  T  S  A  R  L  L  F

 841 ATGGCCGTCAAGTGGGCCAAGAACCTGCCTGTGTTCTCCAGCCTGCCCTTCCGGGATCAG 900
     M  A  V  K  W  A  K  N  L  P  V  F  S  S  L  P  F  R  D  Q
```

FIG.2A

```
901  GTGATCCTGCTGGAAGAGGCGTGGAGTGAACTCTTTCTCCTCGGGGCCATCCAGTGGTCT  960
     V  I  L  L  E  E  A  W  S  E  L  F  L  L  G  A  I  Q  W  S

961  CTGCCTCTGGACAGCTGTCCTCTGCTGGCACCGCCCGAGGCTTCTGCTGCCGGTGGTGCC  1020
     L  P  L  D  S  C  P  L  L  A  P  P  E  A  S  A  A  G  G  A

1021 CAGGGCCGGCTCACGCTGGCCAGCATGGAGACGCGTGTCCTGCAGGAAACTATCTCTCGG  1080
     Q  G  R  L  T  L  A  S  M  E  T  R  V  L  Q  E  T  I  S  R

1081 TTCCGGGCATTGGCGGTGGACCCCACGGAGTTTGCCTGCATGAAGGCCTTGGTCCTCTTC  1140
     F  R  A  L  A  V  D  P  T  E  F  A  C  M  K  A  L  V  L  F

1141 AAGCCAGAGACGCGGGGCCTGAAGGATCCTGAGCACGTAGAGGCCTTGCAGGACCAGTCC  1200
     K  P  E  T  R  G  L  K  D  P  E  H  V  E  A  L  Q  D  Q  S

1201 CAAGTGATGCTGAGCCAGCACAGCAAGGCCCACCACCCCAGCCAGCCCGTGAGGTGACCT  1260
     Q  V  M  L  S  Q  H  S  K  A  H  H  P  S  Q  P  V  R   (SEQ ID NO:2)

1261 GAGCATGCGCCCACCCACTCATCTGTCCCTGACCTCTAACCTTTCTCTGCCTCTCCCACA  1320

1321 CTCTCCCAGAGCTCACTGATTAGACAGCACAAGGGTCTCAGTTCAACAGCATACAGCCAA  1380

1381 CATCTATGGTGTCCCAGGCACAGTGCCAGGCCCCGGGAGTGGGGACCAAGATGTACATAA  1440

1441 GACAAAGCTACTGCCTTCTAGAGACAACCGGCAGTGACCTCACTGAAGACAAAAACTGCC  1500

1501 CTAGCCAGGTACTGAGGGTTGCATGAATCTGCAGGAGACAGAGATCCCCTTGCATGGGAA  1560

1561 ACATAAAGCAGAATTGGGAGGGACTTTGTGGAGACAGGGCTGGACTTGAAAGGAAGAAGA  1620

1621 AGTCTAAAAGAAAACATCATTTGCAAAGGGAGAGAGGGGCAAGCATGATATGTTGTTAGA  1680

1681 ACAGGAGCCCACTTTGAAGGTATAACAGGTTCCTGCCAGTGAGAAATGGGGAGAATAAGC  1740

1741 CAGAAAAGTACCCTAGGACCAGCCCGTTCAGGACTTTGAATGCCAGCCAAAGGCCACGTC  1800

1801 TGACTTGGGAGGCAGAGGGCAGCTACTGCAGGTTTCCGAGCAGAGGGTCATACACAGGGC  1860

1861 TGGACCTCACGCAGACTGGCATGGCCATGGGTCCAGAGGATACTACTGGGAAGGGGATGG  1920

1921 CAGCTACTGCCACCTTCCAGATGGTTCCATGGAGTSCTGATCTTTGGGCATGGCCAGGGG  1980

1981 AAGCAGAAGGGAGACTCTAGGAGTTGAAATGGGTCAGACCCGGTGTTTGGGTGAAGGTAA  2040

2041 GGAATGAGGGAAGAGGAGCTCTTTG      (SEQ ID NO:1)                    2065
```

FIG. 2B

1 METRPTALMS STVAAAAPAA GAASRKESPG RWGLGEDPTG VSPSLQCRVC

51 GDSSSGKHYG IYACNGCSGF FKRSVRRRLI YRCQVGAGMC PVDKAHRNQC

101 QACRLKKCLQ AGMNQDAVQN ERQPRSTAQV HLDSMESNTE SRPESLVAPP

151 APAGRSPRGP TPMSAARALG HHFMASLITA ETCAKLEPED ADENIDVTSN

201 DPEFPSSPYS SSSPCGLDSI HETSARLLFM AVKWAKNLPV FSSLPFRDQV

251 ILLEEAWSEL FLLGAIQWSL PLDSCPLLAP PEASAAGGAQ GRLTLASMET

301 RVLQETISRF RALAVDPTEF ACMKALVLFK PETRGLKDPE HVEALQDQSQ

351 VMLSQHSKAH HPSQPVR (SEQ ID NO:2)

FIG.3

SEQUENCE LISTING

<110> Merck & Co., Inc.

<120> DNA MOLECULES ENCODING HUMAN NUCLEAR RECEPTOR PROTEIN, nNR5

<130> 20083 PCT

<160> 19

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 2065
<212> DNA
<213> Homo sapiens (human)

<400> 1

attcgggacc ttggggcagc tcctgagttc agacagagtt caggaaggga gacaggggca 60
cagagagaca gaggttcatg gactgaggca aaggctgggc caggctcagc aacccaggcc 120
tcccgcaggc aggcagaggc tgccctgtaa cccatggaga ccagaccaac agctctgatg 180
agctccacag tggctgcagc tgcgcctgca gctggggctg cctccaggaa ggagtctcca 240
ggcagatggg gcctggggga ggatcccaca ggcgtgagcc cctcgctcca gtgccgcgtg 300
tgcggagaca gcagcagcgg gaagcactat ggcatctatg cctgcaacgg ctgcagcggc 360
ttcttcaaga ggagcgtacg gcggaggctc atctacaggt gccaggtggg ggcagggatg 420
tgccccgtgg acaaggccca ccgcaaccag tgccaggcct gccggctgaa gaagtgcctg 480
caggcgggga tgaaccagga cgccgtgcag aacgagcgcc agccgcgaag cacagcccag 540
gtccacctgg acagcatgga gtccaacact gagtcccggc cggagtccct ggtggctccc 600
ccggccccgg cagggcgcag cccacggggc cccacaccca tgtctgcagc cagagccctg 660
ggccaccact tcatggccag ccttataaca gctgaaacct gtgctaagct ggagccagag 720
gatgctgatg agaatattga tgtcaccagc aatgaccctg agttcccctc ctctccatac 780
tcctcttcct ccccctgcgg cctggacagc atccatgaga cctcggctcg cctactcttc 840
atggccgtca agtgggccaa gaacctgcct gtgttctcca gcctgccctt ccgggatcag 900
gtgatcctgc tggaagaggc gtggagtgaa ctctttctcc tcggggccat ccagtggtct 960
ctgcctctgg acagctgtcc tctgctggca ccgcccgagg cttctgctgc cggtggtgcc 1020
cagggccggc tcacgctggc cagcatggag acgcgtgtcc tgcaggaaac tatctctcgg 1080
ttccgggcat tggcggtgga ccccacggag tttgcctgca tgaaggcctt ggtcctcttc 1140
aagccagaga cgcggggcct gaaggatcct gagcacgtag aggccttgca ggaccagtcc 1200
caagtgatgc tgagccagca cagcaaggcc caccacccca gccagcccgt gaggtgacct 1260
gagcatgcgc ccacccactc atctgtccct gacctctaac ctttctctgc ctctcccaca 1320
ctctcccaga gctcactgat tagacagcac aagggtctca gttcaacagc atacagccaa 1380
catctatggt gtcccaggca cagtgccagg ccccgggagt ggggaccaag atgtacataa 1440
gacaaagcta ctgccttcta gagacaaccg gcagtgacct cactgaagac aaaaactgcc 1500
ctagccaggt actgagggtt gcatgaatct gcaggagaca gagatcccct tgcatgggaa 1560
acataaagca gaattgggag ggactttgtg gagacagggc tggacttgaa aggaagaaga 1620
agtctaaaag aaaacatcat ttgcaaaggg agagaggggc aagcatgata tgttgttaga 1680
acaggagccc actttgaagg tataacaggt tcctgccagt gagaaatggg gagaataagc 1740
cagaaaagta ccctaggacc agcccgttca ggactttgaa tgccagccaa aggccacgtc 1800
tgacttggga ggcagagggc agctactgca ggtttccgag cagagggtca tacacagggc 1860
tggacctcac gcagactggc atggccatgg gtccagagga tactactggg aaggggatgg 1920
cagctactgc caccttccag atggttccat ggagttctga tctttgggca tggccagggg 1980
aagcagaagg gagactctag gagttgaaat gggtcagacc cggtgtttgg gtgaaggtaa 2040
ggaatgaggg aagaggagct ctttg 2065

<210> 2

<211> 367
<212> PRT
<213> Homo sapien (human)

<400> 2

```
Met Glu Thr Arg Pro Thr Ala Leu Met Ser Ser Thr Val Ala Ala Ala
 1               5                  10                  15
Ala Pro Ala Ala Gly Ala Ala Ser Arg Lys Glu Ser Pro Gly Arg Trp
            20                  25                  30
Gly Leu Gly Glu Asp Pro Thr Gly Val Ser Pro Ser Leu Gln Cys Arg
        35                  40                  45
Val Cys Gly Asp Ser Ser Ser Gly Lys His Tyr Gly Ile Tyr Ala Cys
    50                  55                  60
Asn Gly Cys Ser Gly Phe Phe Lys Arg Ser Val Arg Arg Arg Leu Ile
65                  70                  75                  80
Tyr Arg Cys Gln Val Gly Ala Gly Met Cys Pro Val Asp Lys Ala His
                85                  90                  95
Arg Asn Gln Cys Gln Ala Cys Arg Leu Lys Lys Cys Leu Gln Ala Gly
            100                 105                 110
Met Asn Gln Asp Ala Val Gln Asn Glu Arg Gln Pro Arg Ser Thr Ala
        115                 120                 125
Gln Val His Leu Asp Ser Met Glu Ser Asn Thr Glu Ser Arg Pro Glu
    130                 135                 140
Ser Leu Val Ala Pro Pro Ala Pro Ala Gly Arg Ser Pro Arg Gly Pro
145                 150                 155                 160
Thr Pro Met Ser Ala Ala Arg Ala Leu Gly His His Phe Met Ala Ser
                165                 170                 175
Leu Ile Thr Ala Glu Thr Cys Ala Lys Leu Glu Pro Glu Asp Ala Asp
            180                 185                 190
Glu Asn Ile Asp Val Thr Ser Asn Asp Pro Glu Phe Pro Ser Ser Pro
        195                 200                 205
Tyr Ser Ser Ser Ser Pro Cys Gly Leu Asp Ser Ile His Glu Thr Ser
    210                 215                 220
Ala Arg Leu Leu Phe Met Ala Val Lys Trp Ala Lys Asn Leu Pro Val
225                 230                 235                 240
Phe Ser Ser Leu Pro Phe Arg Asp Gln Val Ile Leu Leu Glu Glu Ala
                245                 250                 255
Trp Ser Glu Leu Phe Leu Leu Gly Ala Ile Gln Trp Ser Leu Pro Leu
            260                 265                 270
Asp Ser Cys Pro Leu Leu Ala Pro Pro Glu Ala Ser Ala Ala Gly Gly
        275                 280                 285
Ala Gln Gly Arg Leu Thr Leu Ala Ser Met Glu Thr Arg Val Leu Gln
    290                 295                 300
Glu Thr Ile Ser Arg Phe Arg Ala Leu Ala Val Asp Pro Thr Glu Phe
305                 310                 315                 320
Ala Cys Met Lys Ala Leu Val Leu Phe Lys Pro Glu Thr Arg Gly Leu
                325                 330                 335
Lys Asp Pro Glu His Val Glu Ala Leu Gln Asp Gln Ser Gln Val Met
            340                 345                 350
Leu Ser Gln His Ser Lys Ala His His Pro Ser Gln Pro Val Arg
        355                 360                 365
```

<210> 3
<211> 860
<212> DNA
<213> Homo sapien (human)

<400> 3
ggaatcacca ggggagacag gngcacagng agacagaggt tcatggactg aggcaaaggc 60
tgggccaggc tcagcaaccc aggcctcccg caggcaggca gaggctgccc tgtaacccat 120
ggagaccaga ccaacagctc tgatgagctc cacagtggct gcagctgcgc ctgcagctgg 180
ggctgcctcc aggaaggagt ctccaggcag atggggcctg ggggaggatc ccacaggcgt 240
gagcccctcg ctccagtgcc gcgtgtgcgg agacagcagc agcgggaagc actatggcat 300
ctatgccctg caacggttgc agcggtttct tccaagagga gcngtacggn ggaggctcaa 360
tccttacaag ggtgcccagg gtgggggcag ggattgtgcc ccccngtgga caaggnccca 420
acccgnaacc cagtgcccag gcctgccggn ttgagaagtg cttnaaaann nggnnggggn 480
ttgaacccag gacgcccgtn naaaggaacg anngccnagc ccgngaggan aagcccaggt 540
nccacccctg ganaagaatn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 600
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 660
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 720
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 780
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 840
nnnnnnnnnn nnnnnnnnnn 860

<210> 4
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 4
atgagctcca cagtggctgc 20

<210> 5
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 5
ctgtctccgc acacgcggca 20

<210> 6
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 6
tttcgagctt ccaggttcat 20

<210> 7
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 7
ctcccaaact ctgcctggtg 20

<210> 8
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 8
cgggagccac acttcaccat 20

<210> 9
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 9
gctcacttct gcgctgtctg 20

<210> 10
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 10
ttccgggctc ccagagtcat 20

<210> 11
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 11
cagaagacct gcctgatctg 20

<210> 12
<211> 20
<212> DNA
<213> Artificial Sequence

<220>

<223> Oligonucleotide

<400> 12
gaaatgaact ccttcatcat 20

<210> 13
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 13
ccggatctgt ggggtgtgtg 20

<210> 14
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 14
ctgatgagaa tattgatgt 19

<210> 15
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 15
cgtgagccgg ccctgggca 19

<210> 16
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 16
ggcatggacc tcactgaaga 20

<210> 17
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Oligonucleotide

<400> 17
actggcagga acctgttata 20

<210> 18
<211> 3012
<212> DNA
<213> Homo sapien (human)

<400> 18
tatagggcga attgggtacc gggccccccc tcgaggtcga cggtatcgat aagcttgata 60
tcgaattcga attcgggacc ttggggcagc tcctgagttc agacagagtt caggaaggga 120
gacaggggca cagagagaca gaggttcatg gactgaggca aaggctgggc caggctcagc 180
aacccaggcc tcccgcaggc aggcagaggc tgccctgtaa ccaatggaga ccagaccaac 240
agctctgatg agctccacag tggctgcagc tgcgcctgca gctggggctg cctccaggaa 300
ggagtctcca ggcagatggg gcctggggga ggatcccaca ggcgtgagcc cctcgctcca 360
gtgccgcgtg tgcggagaca gcagcagcgg gaagcactat ggcatctatg cctgcaacgg 420
ctgcagcggc ttcttcaaga ggagcgtacg gcggaggctc atctacaggt gccaggtggg 480
ggcagggatg tgccccgtgg acaaggccca ccgcaaccag tgccaggcct gccggctgaa 540
gaagtgcctg caggcgggga tgaaccagga cgccgtgcag aacgagcgcc agccgcgaag 600
cacagcccag gtccacctgg acagcatgga gtccaacact gagtcccggc cggagtccct 660
ggtggctccc ccggccccgg cagggcgcag cccacggggc cccacaccca tgtctgcagc 720
cagagccctg ggccaccact tcatggccag ccttataaca gctgaaacct gtgctaagct 780
ggagccagag gatgctgatg agaatattga tgtcaccagc aatgaccctg agttcccctc 840
ctctccatac tcctcttcct ccccctgcgg cctggacagc atccatgaga cctcggctcg 900
cctactcttc atggccgtca agtgggccaa gaacctgcct gtgttctcca gcctgccctt 960
ccgggatcag gtacctaccg gcctgcctgc tggggagcta ggctgggctg gggtcaggcg 1020
gcccactcga gtcaaccaga cagggcacac acatccccac gccagtatga atgcacacag 1080
cttggatggt gatggctggg gacacacata cctctgattc agcgatggct ggggtgcatc 1140
tcagggatgg tgacggtggg ggtgcatgca tctctggcac agggatgatg gtcggggtgc 1200
acacctagga gatgatgatg gctagggacc tacagggccc agggtcttct taagttctgg 1260
aagaccctca ggccctgcag acattctgtg ggtaacaagt gacctgcaca ccctgaacag 1320
gctgagtggc tgactctagg ccccttgga gcacaagtgc ctacgacttc agggcttgca 1380
ttttagttca atctctccag ctctgggcca tccctctcgg cttctaatgg gcaagcagat 1440
ctttcaggaa aaccaggagg agaggcatga ggaaggtttg aggccctcag ccagtctgtg 1500
tgctggggtg gagcaactca gaagagtcag gccacaccac ttgaatacac tcaacttagg 1560
acactcatga ggcatgtctc tgaggctgcc caacttccaa tggctctggg cgttcctaaa 1620
tgtcccagct gcagctctgg atggaaccca gtgtctcaga tgataggcag ctgagccgga 1680
tggtgccaaa tcccagagct ctgagcctct ggctgatgtc aggagagcat tctcgggtcc 1740
caggacagca cttccattcc ttgggtgcct gagatggtgg cagaggctcc agactgagcc 1800
agagaagctg tgtgtctgcc ataacaggca cccctgtctg agcacaggtg atcctgctgg 1860
aagaggcgtg gagtgaactc tttctcctcg gggccatcca gtggtctctg cctctggaca 1920
gctgtcctct gctggcaccg cccgaggcct ctgctgccgg tggtgcccag ggccggctca 1980
cgctggccag catggagacg cgtgtcctgc aggaaactat ctctcggttc cgggcattgg 2040
cggtggaccc cacggagttt gcctgcatga aggccttggt cctcttcaag ccagagacgc 2100
ggggcctgaa ggatcctgag cacgtagagg ccttgcagga ccagtcccaa gtgatgctga 2160
gccagcacag caaggcccac caccccagcc agcccgtgag gtgacctgag catgcgccca 2220
cccactcatc tgtccctgac ctctaacctt tctctgcctc tcccacactc tcccagagct 2280
cactgattag acagcacaag ggtctcagtt caacagcata cagccaacat ctatggtgtc 2340
ccaggcacag tgccaggccc cgggagtggg gaccaagatg tacataagac aaagctactg 2400
ccttctagag acaaccggca gtgacctcac tgaagacaaa aactgcccta gccaggtact 2460
gagggttgca tgaatctgca ggagacagag atccccttgc atgggaaaca taaagcagaa 2520
ttgggaggga ctttgtggag acagggctgg acttgaaagg aagaagaagt ctaaaagaaa 2580
acatcatttg caaagggaga gaggggcaag catgatatgt tgttagaaca ggagcccact 2640
ttgaaggtat aacaggttcc tgccagtgag aaatggggag aataagccag aaaagtaccc 2700
taggaccagc ccgttcagga ctttgaatgc cagccaaagg ccacgtctga cttgggaggc 2760

agagggcagc tactgcaggt ttccgagcag agggtcatac acagggctgg acctcacgca 2820
gactggcatg gccatgggtc cagaggatac tactgggaag gggatggcag ctactgccac 2880
cttccagatg gttccatgga gttctgatct ttgggcatgg ccaggggaag cagaagggag 2940
actctaggag ttgaaatggg tcagacccgg tgtttgggtg aaggtaagga atgagggaag 3000
aggagctctt tg 3012

<210> 19
<211> 2135
<212> DNA
<213> Homo sapien (human)

<400> 19
tatagggcga attgggtacc gggccccccc tcgaggtcga cggtatcgat aagcttgata 60
tcgaattcga attcgggacc ttggggcagc tcctgagttc agacagagtt caggaaggga 120
gacaggggca cagagagaca gaggttcatg gactgaggca aaggctgggc caggctcagc 180
aacccaggcc tcccgcaggc aggcagaggc tgccctgtaa cccatggaga ccagaccaac 240
agctctgatg agctccacag tggctgcagc tgcgcctgca gctggggctg cctccaggaa 300
ggagtctcca ggcagatggg gcctggggga ggatcccaca ggcgtgagcc cctcgctcca 360
gtgccgcgtg tgcggagaca gcagcagcgg gaagcactat ggcatctatg cctgcaacgg 420
ctgcagcggc ttcttcaaga ggagcgtacg gcggaggctc atctacaggt gccaggtggg 480
ggcagggatg tgccccgtgg acaaggccca ccgcaaccag tgccaggcct gccggctgaa 540
gaagtgcctg caggcgggga tgaaccagga cgccgtgcag aacgagcgcc agccgcgaag 600
cacagcccag gtccacctgg acagcatgga gtccaacact gagtcccggc cggagtccct 660
ggtggctccc ccggccccgg cagggcgcag cccacggggc cccacaccca tgtctgcagc 720
cagagccctg ggccaccact tcatggccag ccttataaca gctgaaacct gtgctaagct 780
ggagccagag gatgctgatg agaatattga tgtcaccagc aatgaccctg agttcccctc 840
ctctccatac tcctcttcct ccccctgcgg cctggacagc atccatgaga cctcggctcg 900
cctactcttc atggccgtca agtgggccaa gaacctgcct gtgttctcca gcctgccctt 960
ccgggatcag gtgatcctgc tggaagaggc gtggagtgaa ctctttctcc tcggggccat 1020
ccagtggtct ctgcctctgg acagctgtcc tctgctggca ccgcccgagg cctctgctgc 1080
cggtggtgcc cagggccggc tcacgctggc cagcatggag acgcgtgtcc tgcaggaaac 1140
tatctctcgg ttccgggcat tggcggtgga ccccacggag tttgcctgca tgaaggcctt 1200
ggtcctcttc aagccagaga cgcggggcct gaaggatcct gagcacgtag aggccttgca 1260
ggaccagtcc caagtgatgc tgagccagca cagcaaggcc caccacccca gccagcccgt 1320
gaggtgacct gagcatgcgc ccacccactc atctgtccct gacctctaac ctttctctgc 1380
ctctcccaca ctctcccaga gctcactgat tagacagcac aagggtctca gttcaacagc 1440
atacagccaa catctatggt gtcccaggca cagtgccagg ccccgggagt ggggaccaag 1500
atgtacataa gacaaagcta ctgccttcta gagacaaccg gcagtgacct cactgaagac 1560
aaaaactgcc ctagccaggt actgagggtt gcatgaatct gcaggagaca gagatcccct 1620
tgcatgggaa acataaagca gaattgggag ggactttgtg gagacagggc tggacttgaa 1680
aggaagaaga agtctaaaag aaaacatcat ttgcaaaggg agagaggggc aagcatgata 1740
tgttgttaga acaggagccc actttgaagg tataacaggt tcctgccagt gagaaatggg 1800
gagaataagc cagaaaagta ccctaggacc agcccgttca ggactttgaa tgccagccaa 1860
aggccacgtc tgacttggga ggcagagggc agctactgca ggtttccgag cagagggtca 1920
tacacagggc tggacctcac gcagactggc atggccatgg gtccagagga tactactggg 1980
aaggggatgg cagctactgc caccttccag atggttccat ggagttctga tctttgggca 2040
tggccagggg aagcagaagg gagactctag gagttgaaat gggtcagacc cggtgtttgg 2100
gtgaaggtaa ggaatgaggg aagaggagct ctttg 2135

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US98/26422

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :C07K 14/435, 14/705; C12N 5/10, 15/10, 15/12, 15/63
US CL :435/69.1, 320.1, 325; 530/350; 536/23.5
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/69.1, 320.1, 325; 530/350; 536/23.5

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, STN, BIOSCIENCE, MEDLINE, CAPLUS, BIOSIS, GENBANK, PIR, EST

search terms: nuclear receptor, steroid receptor, nNR5

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	MANGELSDORF et al. The RXR heterodimers and orphan receptors. Cell. 15 December 1995, Vol. 83, pages 841-850.	1-29

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 19 MARCH 1999	Date of mailing of the international search report 15 APR 1999
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer MICHAEL PAK Telephone No. (703) 308-0196